An investigation of the spatial patterning of gambling-related harm and the total consumption theory of gambling

Francis Markham

April 2017

A thesis submitted for the degree of Doctor of Philosophy
The Australian National University

Australian National University

Candidate's declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of the author's knowledge, it contains no material previously published or written by another person, except where due reference is made in the text.

Francis Markham

April 2017

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisors Bruce Doran and Martin Young. It is no exaggeration to say that none of this work would have been possible without their guidance, patience, support and good humour. I am deeply grateful for the huge amount of time and energy they have both dedicated to my training over many years. It has truly been a privilege to be supervised by them. I could not have hoped for a better supervisory team.

I am indebted to the people who have generously participated in the surveys analysed in this thesis. All have given a considerable amount of time to complete questionnaires or engage in interviews. Many have reported experiences of considerable suffering. This research is dedicated to them.

I am also grateful to several senior colleagues in Australian and British gambling studies who – although not directly connected to my thesis – reassured me during the course of this research. Gambling research is, at times, a bewildering and disheartening field of study for a new scholar to navigate, beset by conflicts of interest and filled with intimidating personalities. I owe a debt of gratitude to those whose encouragement and counsel helped me to steer a course through these sometimes-difficult waters.

This thesis has relied on data collected by others. I am grateful to the Community Benefit Fund of the Northern Territory Government, the Australian Research Council and the Northern Territory Research and Innovation Board who jointly funded the collection of the NT survey data analysed in Chapters 2, 4 and 6 of this thesis. The Northern Territory Department of Justice provided the EGM venue data analysed in these same chapters. For the survey data that was reanalysed in Chapter 3, I am grateful to many groups. The Australian National Gambling Survey was conducted by the Productivity Commission and funded by an appropriation from the Treasury of the Commonwealth of Australia. The Canadian National Validation Survey was conducted by the Canadian Centre on Substance Abuse and was funded by Canadian provincial governments. The Finnish Gambling Survey 2011 was conducted for the National Institute for Health and Welfare and funded by the Ministry of Social Affairs and Health of Finland. The Gambling in Norway survey was funded by the Norwegian Institute for Alcohol and Drug Research (SIRUS) and the Norwegian Lottery Board. I am grateful to the Australian Data Archive, the Gambling Research Exchange Ontario, the Finnish Social Science Data Archive and Ingeborg Lund of the Norwegian Institute for Alcohol and Drug Research for making these data available. I am grateful to Paul Delfabbro for his assistance in extracting data from an unpublished report analysed in Chapter 5. Finally, Chapter 7 would not have been possible without the provision of data by Crime Statistics Agency of Victoria and the Victorian Commission for Gambling and Liquor Regulation.

This thesis would not have been possible without an Australian Postgraduate Award from the Commonwealth Government. This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

Many scholars provided valuable feedback on the work compiled in this thesis. Most importantly, anonymous reviewers provided detailed critiques of earlier versions of the manuscripts published in this thesis which substantially improved the work. Bryan Rodgers and Ingeborg Lund both offered helpful suggestions regarding the work presented in Chapter 3. Sandra Potter and Nicola Sloan also offered useful advice about several of the studies collected here. I am particularly grateful to Mishka Talent for statistical advice and for his insistent yet constructive scepticism toward social research methods, which challenged me to approach several of the methods and metrics used here more critically. Members of the Health Geography Study Group of the Institute of Australian Geographers provided helpful comments on Chapter 7.

I am grateful to everyone at the Fenner School of Environment and Society for providing a collegial, friendly and relaxed study environment. I'd particularly like to thank Amy Chen and Jack Pezzey for making administrative processes as smooth as possible. Thanks also to my fellow students – especially Sandra Potter, Mishka Talent, Mae Noble, Claudia Benham, Richard Beggs, Kat Ng, Diane Erceg, Boas Bangita and many others – for the chats, company, coffees and good times. Thanks also to Inger Mewburn and Cathy Ayres at ANU Research Training for their support.

Last and most important, I would like to express my deep appreciation to Hannah Quinlivan. Thank you for your encouragement, care and support, for delighting in my successes and for commiserating with me in more frustrating moments. And thank you for companionably sharing the weekends and long evenings devoted to this thesis, even as we have worked on our separate projects. Thanks to Chris Higgisson and Sally Markham for supporting this project in a thousand ways. This thesis would not have come about without your encouragement and care.

Declaration of competing interests

Francis Markham has received funding from the Alberta Gambling Research Institute (AGRI), a consortium of researchers at three universities, to present a keynote at AGRI's annual conference. AGRI is funded by an appropriation from the Alberta Department of Treasury Board and Finance, and is governed by an eight-member board, seven of whom are faculty members of universities in Alberta, and one of whom represents the public and is elected by the other board members. AGRI's appropriation from the Department of Treasury Board and Finance is voted on by the provincial government of Alberta as part of the departments' supply vote, and is drawn from the Alberta Lottery Fund. The Alberta Lottery Fund receives and disburses profits from the province's state-owned gambling monopolies, including those on video lottery terminals, slot machines and lottery tickets.

Markham has been employed on research projects funded by the Australian Capital Territory Gambling and Racing Commission, and has been employed on a project jointly funded by the Australian Research Council, the Community Benefit Fund (Northern Territory Government), and the Northern Territory Research and Innovation Fund (Northern Territory Government). The Community Benefit Fund is a government administered grant scheme, funded by a levy on EGMs in hotels in the Northern Territory.

Markham is a member of the Public Health Association of Australia (PHAA) and is the ACT representative of the PHAA Political Economy of Health Special Interest Group.

Abstract

Gambling is an important public health issue in Australia. According to recent estimates, gambling-related harm is the third largest contributor to the burden of disability in the state of Victoria, measured in terms of disability-adjusted life years. The gambling product most associated with gambling-related harm in Australia is the electronic gaming machine (EGM), which accounts for over half of all Australian gambling expenditure. Around 30 per cent of weekly EGM gamblers experience moderate or severe adverse impacts from their gambling.

This thesis consists of six studies on the spatial distribution of the impacts of electronic gaming machines (EGMs) and the relationship between EGM losses and problem gambling. All have been published or were accepted for publication in peer-reviewed academic journals at the time of submission.

Jointly, these studies developed theoretical and methodological tools to advance the production of small area estimates of gambling-related harm, as well as beginning the exploration of its consequences. The six studies in this thesis can be grouped into three inter-linked themes that contribute to this aim in different ways.

Two studies are concerned with developing the applied and methodological tools for investigating the spatial distribution of problem gambling. The first of these studies presents a calibrated Huff model of the spatial behaviour of gamblers. The second of these uses the Huff model to refine spatial microsimulation derived small area estimates of the prevalence of problem gambling. Together, they provide a toolkit for estimating the local impacts of EGMs.

Three studies provide the theoretical underpinning of the thesis by investigating the relationship between gambling losses and problem gambling at the scales of the individual, the EGM venue and state or territory. In order to develop the methods for investigating the spatial distribution of problem gambling, a sustained engagement was required with Total Consumption Theory in the context of gambling. These studies find a consistent relationship between EGM losses and the risk of harm at all spatial scales. At the scale of the individual, there is no evidence to support a J-shaped dose-response

relationship, meaning that risk of gambling problems increases monotonically with money lost.

A final study estimates the spatio-temporal correlation between EGM accessibility and a single gambling-related harm, domestic violence. Whereas research in the earlier phases of this project sought to estimate the distribution of ‘problem gambling’ as an outcome measure, phase four seeks to measure the relationship between EGM accessibility and specific gambling-related harms directly. In this instance, the spatial association between EGMs and police-recorded domestic violence incidents is investigated in Victorian postcodes over a ten-year period. A significant spatio-temporal association between these two variables is found, providing evidence of a link between EGM gambling and violence. This study concludes that future research might usefully explore the spatio-temporal co-occurrence of EGM gambling and specific gambling-related harms to better understand the social and health impacts of EGM gambling.

The research developed in this thesis has contributed toward bringing knowledge of the geography of the impacts of EGMs closer to that of cognate public health issues. While Total Consumption Theory was developed in the context of gambling to underpin the production of local area estimates that incorporate gambling consumption as a risk factor, the findings in this section have broader implications for gambling regulation. More broadly, the approaches developed in this thesis and the research findings have the potential to contribute to improving the regulation of EGMs and thereby reduce the incidence of gambling-related harms.

Table of contents

List of tables

List of figures

Chapter 1: Context and research approach

1.1 Background

Electronic gaming machines (EGMs) proliferated across Australia in the 1990s. Australian EGMs are a high-intensity variant of Las Vegas-style slot machines, known in the vernacular as 'poker machines' or 'pokies'.[1] The increase in EGM accessibility and EGM gambling losses led to an increasing awareness of gambling-related harm and problem gambling among the public, policy makers and regulatory bodies. This thesis contributes to the study of the social and health impacts of EGMs in Australia. It explores these impacts using an explicitly spatial approach. Specifically, it makes three key contributions to knowledge:

1. It refines geographic methods for understanding the spatial behaviour of EGM gamblers and estimating the impacts of EGM gambling in small geographic areas

2. It clarifies the relationship between EGM gambling losses and gambling impacts for individuals, gambling venues and jurisdictions

3. It commences a new research agenda focussed on estimating the spatial relationship between EGM consumption and specific gambling-related harms

1.1.1 The proliferation of electronic gaming machines in Australia

Australia was among the first countries to legalise mechanical gaming machines, with clubs in New South Wales permitted to operate the machines in 1956 (Australian Institute for Gambling Research, 1999; Chambers, 2011). Yet EGMs were not widely legalised in Australia until the 1990s. In 1989 EGMs were banned from hotels and clubs in every state and territory of Australia except New South Wales and the Australian Capital Territory. A wave of legalisation in the 1990s had the result that by the decade's close, EGMs were accessible in hotels and clubs in every jurisdiction except Western

[1] Australian EGMs are a high-intensity gambling machine. Gambler can bet on multiple lines simultaneously on Australian EGMs, and bets can be 'multiplied' several times. They generally feature large jackpot prizes. Typical losses on these machines when played at high intensity ranges from around \$600 to \$1200 per hour (Productivity Commission, 2010). These are distinct from lower intensity gaming machines, such as British fruit machines, Japanese pachinko machines or Canadian Video Lottery Terminals (although some 'multigame' video lottery terminals do include multiline slot machine games) (MacLaren, 2015; Schüll, 2012; Turner and Horbay, 2004).

Australia (Australian Institute for Gambling Research, 1999). In 1999 there were 184,526 EGMs operating in Australia, or one for every 76 adults (Productivity Commission, 1999).

The increased accessibility of EGMs led to a dramatic increase in gambling losses in Australia. As Figure 1.1 shows, variation in total gambling losses in Australia in each state and territory is largely driven by changing losses on EGMs. Whereas in 1989-90 per capita EGM loss in hotels and clubs was just $217 (in 2015 dollars), by 2001-02 this figure had increased to $861 (Queensland Government Statistician's Office, 2016). As a percentage of total gambling losses, spending on EGMs in hotels and clubs rose from 32% of total losses in 1989-90 to 60% of gambling losses in 2001-02.

Increased public concern accompanied the growing accessibility of, and losses on, EGMs in Australia. For example, by 1998, six years after EGMs were legalised in Victoria, 82% of adults agreed that 'gambling-related problems have got worse in the last 4 years' and 68% wanted to see the number of EGMs in the state decreased (Roy Morgan Research, 1999). In the state of South Australia where EGMs were introduced in 1994, the 1997 election saw an independent candidate elected to parliament campaigning on a 'no pokies' platform (McCarthy, 1999). In response to growing community concern, the federal government of Australia requested that the leading government research body, the Productivity Commission, inquire into Australia's gambling industries in order to better understand their growing social and economic impacts. The Productivity Commission's resultant (1999) report established the status of EGM gambling as a significant social and health issue in Australia, and set much of the scholarly research agenda on gambling over subsequent decades.

Following the Productivity Commission's (1999) report, each Australian jurisdiction introduced two key reforms which have been credited with reducing per capita EGM expenditure: caps on EGM numbers and smoking bans (Productivity Commission, 2010). EGM caps have been introduced using different mechanisms in different states and territories, including caps on the number of EGMs per venue, regional and municipal caps on the number of EGMs in particular areas and caps on the total number of EGMs in an entire state or territory (Productivity Commission, 2010). A consequence of the system of caps has been a fall in EGM accessibility. The number of

EGMs in Australian hotels and clubs peaked at 200,507 in 2004-05 and gradually declined to 196,661 in 2014-15 (Queensland Government Statistician's Office, 2016). Once population growth is accounted for, EGM density fell from 131 EGMs per 10,000 adults in 2004-05 to 107 EGMs per 10,000 adults in 2014-15.

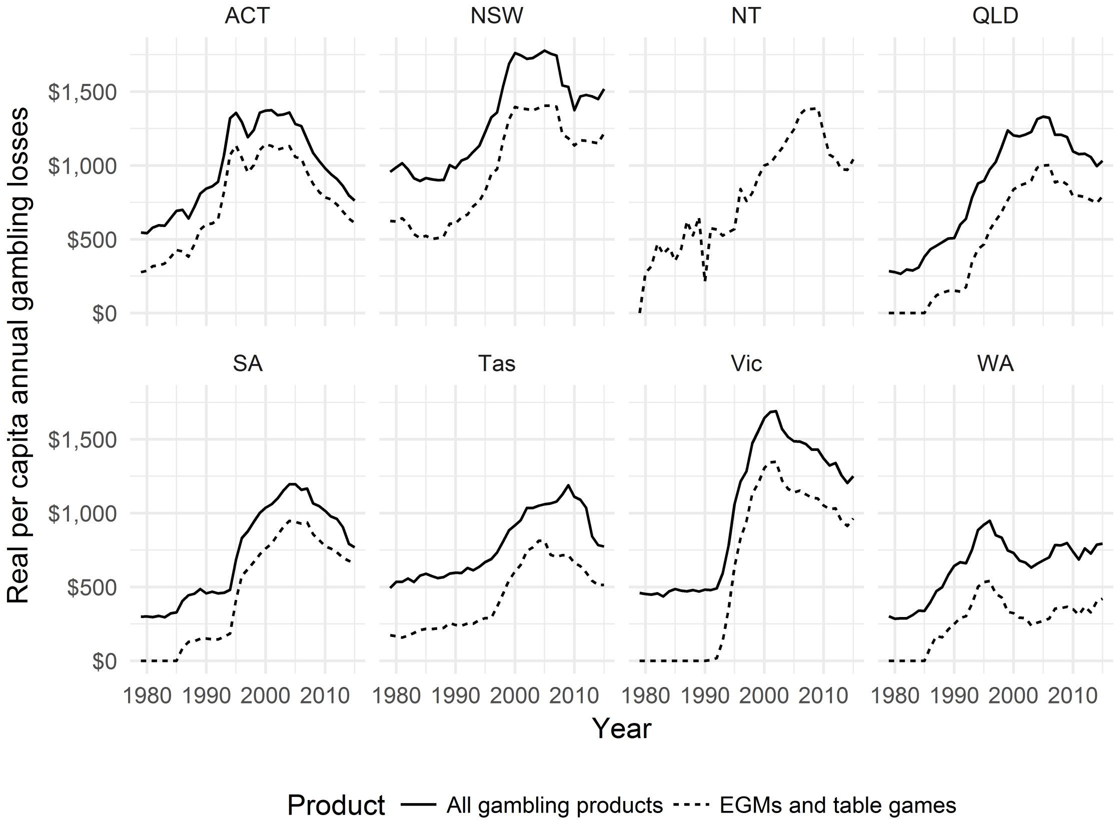

Figure 1.1: Real per capita player losses on EGMs and casino table games, and other gambling products, 1978-79 to 2014-15. EGMs and table games are aggregated because player losses in casinos are not disaggregated by product. Source: Queensland Government Statistician's Office (2016). Total gambling losses for the Northern Territory are excluded due to exaggeration resulting from national losses on all major corporate internet bookmakers being reported as losses located in the Northern Territory.

Smoking bans in EGM venues have also contributed to a reduction in per capita EGM gambling losses. Between 2002 and 2010 Australian states and territories all introduced indoor smoking bans in EGM venues as a tobacco control measure (Livingstone et al., 2014). Although they are yet to be evaluated nationally, the introduction of 'smoke-free' policies in Victoria and South Australia was associated with reduction in EGM gambling losses of approximately 14% (Hirschberg and Lye, 2010; Lal and Siahpush, 2008). The effectiveness of smoking-free policies may have been attenuated in New

South Wales as outdoor EGM gambling areas were permitted in that state after an indoor smoking ban was introduced (Livingstone et al., 2014).

The effect of these and other measures in reducing gambling losses has been underpinned by falling participation in EGM gambling. A compilation of estimates from state-based surveys suggests that in all states and territories, the proportion of the population gambling on EGMs at least once per year peaked between 2000 and 2002 (see Figure 1.2). The process whereby a decreasing proportion of the population use EGMs over time has been prominently described as 'adaptation' (Abbott, 2006), although it is unclear to what extent adaptation has been driven by changing regulations and legislation, changing social attitudes or difficulties in conducting comparable surveys in an environment characterised by falling survey response rates and a failure to contact the growing subpopulation that is not reachable by random digit dial surveys of fixed-line telephones (Markham and Young, 2016).

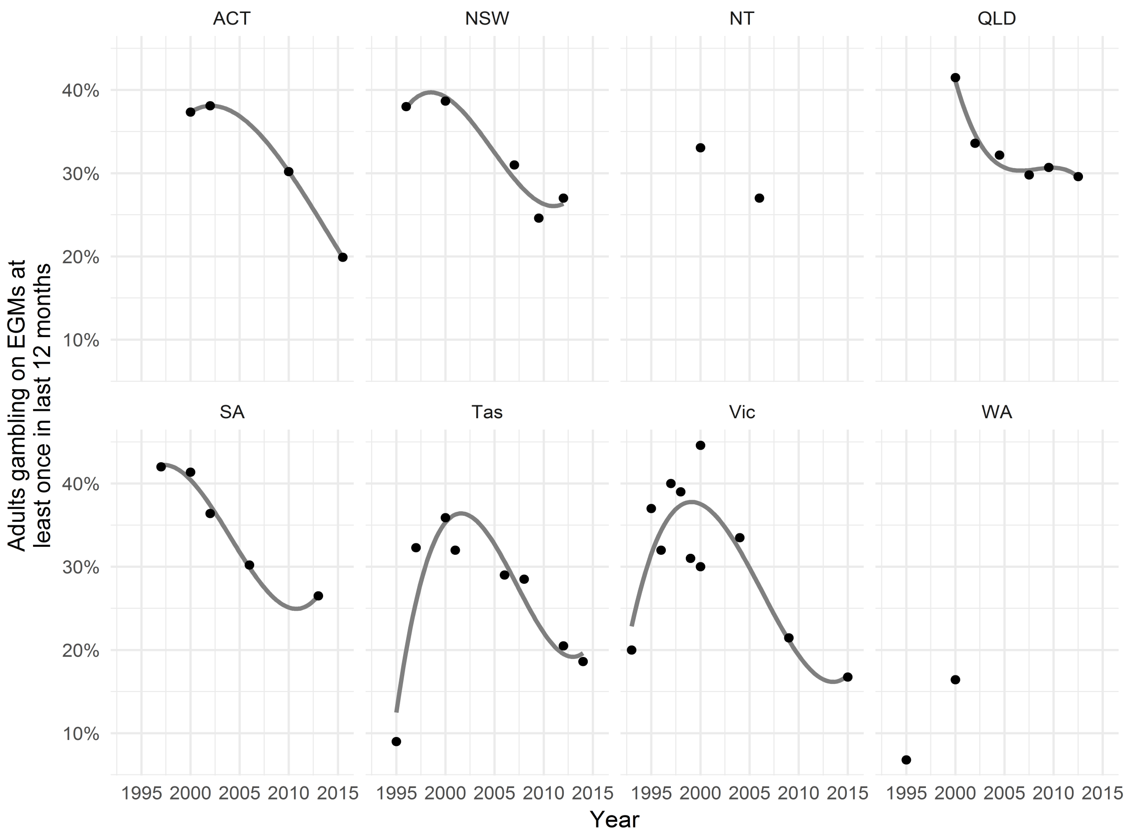

Figure 1.2: Estimated annual EGM participation in Australian jurisdictions, 1994-2015. Points represent individual surveys. Curved lines show the regression fit from third-order polynomial regressions. Source: Author's compilation from 43 published survey estimates. Comparisons of estimates should be with caution due to substantial differences in methodology among surveys.

Despite the introduction of systems of caps on EGM numbers and other regulatory interventions, Australia retains an unusually high level of EGM accessibility in hotels and clubs. Cross-national comparisons show that Australia has the highest density of EGMs of any country in the world (excluding resort destinations such as Monaco) (Ziolkowski, 2016), with the density of EGMs outside casinos especially high (Chambers, 2011). Consequently, Australians lose far more per capita gambling than residents of any other country in the world. As Figure 1.3 shows, Australia's exceptionalism is driven by the unequalled level of gambling losses on EGMs outside casinos in Australia.

Two reasons for Australia's exceptionalism may be given. A distal cause relates to Australia's overarching regulatory regime. Chambers (2011) links gambling regulation cross-nationally with the type of welfare regime. He suggests that the availability of EGMs in Australia is a result, in part, of Australia's liberal state apparatus, which prioritises a residual safety net social welfare and minimal state intervention in the market. While this explains the limited regulation of EGMs in Australia, it does little to explicate why other states with similarly liberal regimes do not have similarly high EGM expenditure.

Proximately, while many other countries have high rates of gambling, Australia is unique in that it's states and territories facilitate the provision of large numbers of high-intensity EGMs throughout the urban fabric. In Singapore and Ireland, the second and third ranked countries in terms of per capita gambling expenditure, there is little access to EGMs. In Singapore, EGMs are restricted to two casinos only, which require an identification check on entry and charge a substantial entry levy to Singapore residents (Hancock and Hao, 2016). In Ireland, EGMs are banned entirely by statute. In Australia, however, EGMs have been legal since 1956 and the entrenched nature of the industry ensures that efforts at regulation meet considerable resistance from interest groups (Markham and Young, 2015).

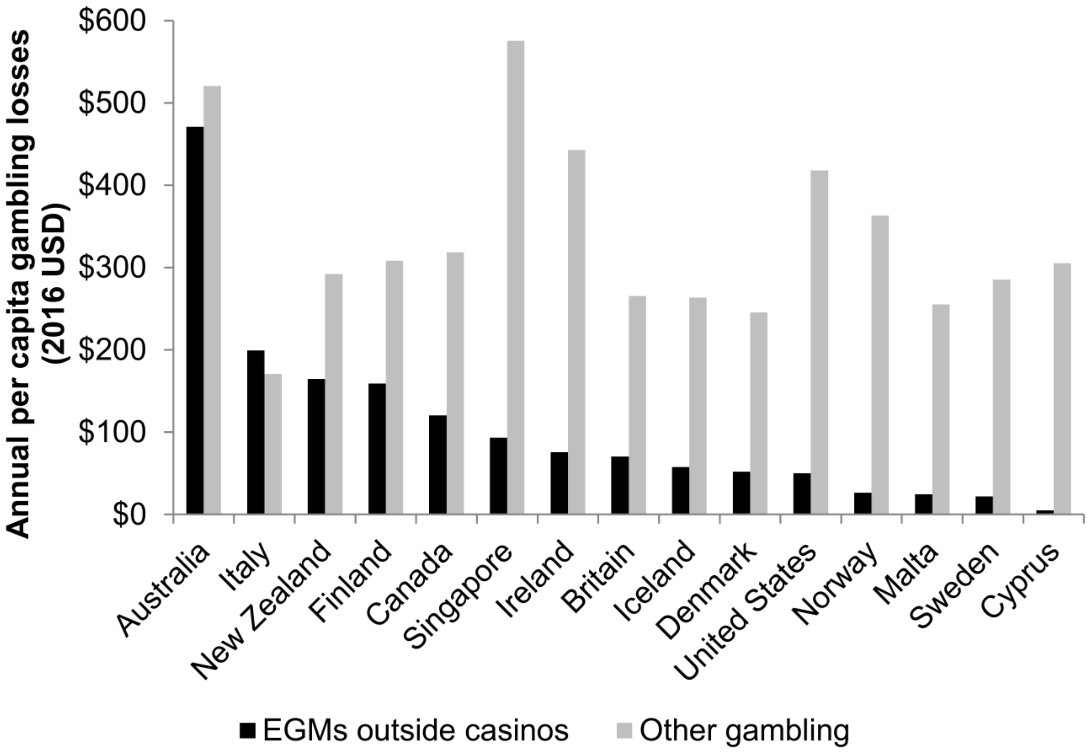

Figure 1.3: Per capita annual gambling losses on EGMs outside casinos and on other activities in 2016 for the fifteen highest per capita spending countries. Source: H2 Gambling Capital, cited by The Economist (2017).

1.1.2 Gambling on electronic gaming machines as a public health issue

The proliferation of EGM gambling is a significant public health concern because a substantial proportion of people who gamble experience adverse consequences from their gambling. For example, in the most recent nationally-representative population survey, 5.9% of those who gambled in the previous twelve months reported sometimes betting more than they could afford, and 3.4% felt that their gambling had caused them health problems (Dowling et al., 2016). For individuals, increases in the frequency of gambling and the amount lost are associated with increased levels of harm (Productivity Commission, 2010, 1999). This suggests that gambling should be considered as a public health 'exposure' which puts individuals at risk of experiencing adverse consequences (Rodgers et al., 2009).

There is no universally accepted definition of what constitutes a gambling-related harm. However, more progress has been made in describing and classifying the range of harms that gamblers report experiencing. The most comprehensive effort of this sort suggested that gambling-related harms for individuals range across seven domains (Langham et al., 2016):

- financial harms (the loss of discretionary spending, the inability to pay for household expenses, the accrual of debts, and the loss of assets such as houses)

- harms to relationships (the loss of time spent with others, relationship conflicts, and relationship breakdowns such as separations or divorce)

- emotional or psychological harms (feelings of a loss of control, reduced sense of security or safety, and shame and stigma)

- health impacts (increased drinking and smoking, reduced exercise and sleep, depression and anxiety, inability to afford healthcare, increased stress and blood pressure, violence including family violence, self-harm, and suicide)

- impacts on work or study (increased absenteeism, reduced performance, and termination of employment or study due to poor performance)

- criminal acts (negligence such as child neglect, and crimes committed to gain access to funds to gamble with or to repay debts such as drug trafficking, petty theft from family members, illicit lending, and fraudulent efforts to attain funds)

- cultural harms (the loss of time spent on cultural practices and roles, dissonance between gambling and cultural beliefs, and feeling of lost cultural identify)

In addition, gambling has negative social and economic impacts that are felt at neighbourhood, municipal and jurisdictional scales, although these are rarely the focus of gambling research (Productivity Commission, 2010).

These gambling-related harm translate into adverse health outcomes via a variety of pathways, most of them indirect. Directly, there are several clear pathways. Gambling increases the chances of comorbidity with other forms of health-adverse addictive consumption, such as drinking alcohol and smoking (Peters et al., 2015). Excessive gambling also leads directly to mental health problems, such as major depression and anxiety, as a direct consequence of difficulties controlling gambling and money and time lost gambling (Yakovenko & Hodgins, 2018). Longitudinal studies also suggest that stress related to gambling results in an increased incidence of some heart conditions in older adults (e.g. Pilver & Potenza, 2013).

Indirectly, and perhaps more importantly, gambling is likely to impact on health by effecting the social determinants of health. A vast literature now exists documenting how economic resources and position on the social gradient translate into poorer health outcomes (see, for example, Marmot, 2005). By requiring the expenditure of a great deal of time and money, gambling increases the social disadvantage of gamblers and their families. It is likely that the bulk of the harms arising from gambling are transmitted through the vector of reduced economic resources, and the concomitant flow on effects that this has for gamblers and their families across the life course. However, research into the detail of these mechanisms and their efficacy is immature at this time.

Young (2013) argues that the social and health impacts of gambling are poorly understood, in part because the field has until recently been characterised by a focus on an alternative conceptualisation to gambling-related harm, that of ‘problem gambling’. According to the Australian national definition, problem gambling is “…characterised by difficulties in limiting money and/or time spent on gambling which leads to adverse consequences for the gambler, others, or for the community” (Neal et al., 2005, p. 124). The constructs of ‘problem gambling’ and ‘problem gamblers’ are dominant in the gambling research literature (Cosgrave, 2010; Miller et al., 2016; Reith, 2007; Young, 2013). While the national definition of problem gambling emphasises adverse consequences, the psychometric instruments for identifying problem gambling in the research literature are based around identifying individual problem gamblers on the basis of their displaying symptoms of addiction or an impulse control disorder (Svetieva and Walker, 2008). In practice, people tend to be classified in the research literature as a ‘problem gambler’ if they report clinical or subclinical levels of the psychologically-defined conditions of pathological or disordered gambling (Productivity Commission, 2010). Put simply, problem gambling should be understood as a widely used alternative construct to gambling-related harm which is focussed on gamblers’ psychological states rather than the adverse impacts of their gambling. Much of the research in this thesis utilises this, admittedly limited, conceptualisation of gambling-related harm due to issues of data availability and the lack of any validated instrument for measuring gambling-related harm to date.

The impact of gambling-related harms in Australia is vast in scale. Quantified in terms of money, it has been conservatively estimated that the social costs of problem gambling in Australia amount to between $4.7 and $8.4 billion per annum (Productivity Commission, 2010). The social cost of gambling is comparable to that of alcohol, which was most recently estimated to be $15.3 billion per annum (Collins and Lapsley, 2008). More recently, the 'burden of disease' attributable to gambling has been estimated in the state of Victoria using the standard metric of quality adjusted life years (Browne et al., 2016). This study found that, at the population level, the burden of disease attributable to gambling was 69% of that attributable to alcohol use and dependence, 205% of that attributable to osteoarthritis and 447% of that attributable to diabetes mellitus. While these estimates are admittedly inexact, it is clear that gambling has considerable public health impacts. Although the conceptualisation of the negative impacts of gambling has evolved over several decades (from 'problem gambling' to 'gambling-related harm') and methods for estimating the magnitude and prevalence of these impacts have improved, the social impacts of gambling are unmistakably extensive.

The negative impacts of gambling are most acute in the case of EGMs (Productivity Commission, 2010). The impact of different gambling products can be measured in many different ways, but all of these converge to suggest that EGMs are the most problematic. For example, around 80% of those who seek treatment for gambling problems report that their difficulties primarily result from EGM gambling (Productivity Commission, 2010). EGM revenues most heavily accrue from 'problem gamblers' (Rodgers et al., 2015), with between 22% and 60% of all EGM losses originating from the 0.7% of the population so classified (Productivity Commission, 2010). Among the 4% of the Australian population who play EGMs weekly, 20-45% are classified as being either problem gamblers or being at 'moderate risk' of developing gambling problems, the highest proportion for any gambling product (Productivity Commission, 2010). Finally, the majority of money lost gambling in

Australia is lost on EGMs (Queensland Government Statistician's Office, 2016), making their especially harmful nature a significant public health concern.

1.1.3 The spatial distribution of gambling-related harm

Relatively little is known about the spatial distribution of gambling-related harm. Research with a geographic focus has primarily investigated three sets of relationships (St-Pierre et al., 2014; Vasiliadis et al., 2013):

1. Socio-economic status and EGM density

2. Density of EGMs and gambling behaviour (including gambling expenditure and problem gambling)

3. Residential proximity to EGM venues and gambling behaviour (including problem gambling)

Research on socio-economic status and EGM density has consistently found that EGMs are spatially concentrated in the poorest areas of cities, both in Australia and elsewhere. In a series of studies pioneered in Adelaide and replicated in Sydney, Melbourne and the Richmond-Tweed, Marshall documented a highly consistent relationship between socio-economic status and EGM density (Marshall, 2005, 1999; Marshall and Baker, 2001a, 2001b, 2000). Perhaps most crucially, Marshall and Baker (2002) demonstrated that this relationship is no historical accident but is the result of a systematic process of socio-spatial allocation. In particular, Marshall and Baker (2002) show how, at the time of legalisation, EGMs were randomly distributed across Victorian local government areas, before becoming increasingly concentrated in poorer areas over time.

The allocation of EGMs to poorer areas is of importance because the spatial distribution of EGM density has been linked to the spatial distribution of gambling-related harm. This finding was first documented by the Productivity Commission (1999) report which demonstrated both that gambling expenditure was higher in local government areas with a greater EGM density and that problem gambling prevalence was higher in states and territories with a greater EGM density. The relationship between EGM density and problem gambling was confirmed by Storer *et al.* (2009), whose analysis of the results of 34 prevalence studies found a linear association between EGM density and problem gambling prevalence, with each EGM in a jurisdiction associated with 0.8 problem gamblers.

Yet these associational studies say little about the nature of the spatial relationship involved. As the Productivity Commission noted (1999), the spatial distribution of EGMs within a given spatial unit may be as important as EGM density *per se*. Using an example of two cities with the same number of EGMs but different spatial configurations of EGM venues (Figure 1.4), the commission raised the question of the extent to which the EGM density – problem gambling nexus was mediated by accessibility.

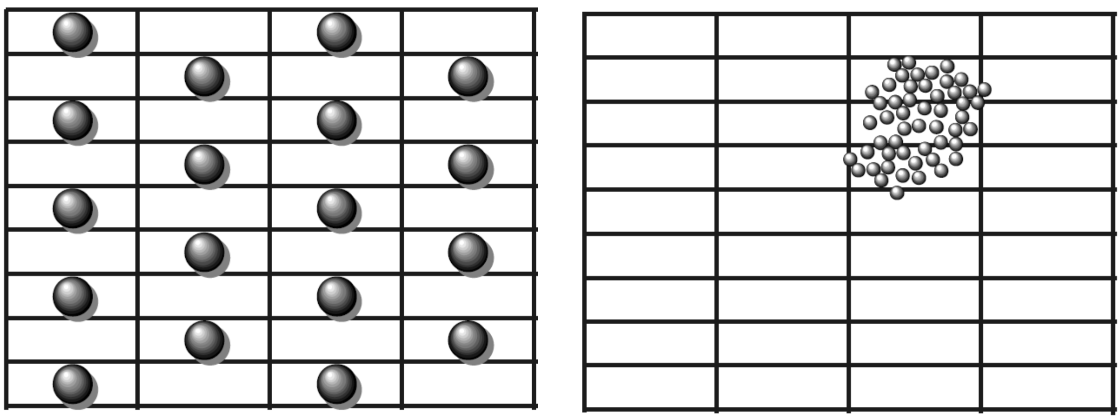

Figure 1.4: 'Does spatial distribution affect accessibility? Two cases'. Source: Productivity Commission (1999, p. 8.5). In the left panel, a small number of large EGM venues are evenly spaced across a hypothetical city. In the right panel, a large number of small EGM venues are spatially clustered in one area of a hypothetical city.

Two strands of research have sought to answer this question. The first sought to evaluate the distance travelled to EGM venues in actual cities. Perhaps the best example of this approach was a geocoded door-knock survey of suburban Tuggeranong, Australia, which asked people which venues they visited and connected that to their home address (Doran et al., 2007). This study found that while some venues had highly localised catchments, other venues had much more spatially extensive catchment areas, particularly those located near shopping centres. This finding implies that a straightforward allocation of EGMs' impacts to the containing administrative unit may not accurately capture actual patterns of spatial behaviour. Further, it showed that the decision to visit a particular venue is multifactorial, and dependent on factors beyond spatial accessibility. What has been called 'social accessibility,' for example community ties to a particular venue, appears to also play a role in determining which EGM venues gamblers visit (Productivity Commission, 1999).

The second strand of research examined proximity to EGM venues in addition to EGM density. In one example of this research, Young *et al.* (2012a) found that residential proximity to EGM venues increased the probability of gambling participation and increased gambling frequency, thereby increasing problem gambling risk. Similar findings have been reported internationally (e.g. Pearce et al., 2008; Rush et al., 2007; John W Welte et al., 2004). The national geocoded survey analysed by Pearce *et al.* (2008) is particularly suggestive of the importance of proximity rather than density as the key spatial relationship driving the association between accessibility and harm. This study, which evaluated the association between problem gambling and both EGM density and EGM proximity found that once proximity is accounted for, density is no longer predictive of problem gambling. In other words, EGM venues located beyond the venue that is nearest to home did not appear to influence problem gambling risk.

One key limitation of the majority of the studies described above is the ecological fallacy, deriving in large part from the scale of geographical analysis. The ecological fallacy can occur when the relationship of two variables at one spatial scale differs from the relationship between the same variables at a smaller spatial scale (Haining, 2003). Specifically, Australian studies have typically analysed Local Government Areas (LGAs: e.g. Allen Consulting Group, 2012; Marshall, 1999; Marshall and Baker, 2002, 2000; Productivity Commission, 1999), an administrative unit with a median population of around 11,000 people (*IQR* = 2587, 39356) and a median extent of 2,371 km^2 (*IQR* = 321, 5972 km^2). Studies that have examined scales smaller than the LGA have typically reported more equivocal findings (e.g. Marshall and Baker, 2001a), especially with regard to the relationship between disadvantage and expenditure (McMillen and Doran, 2006). As Doran, Marshall and McMillen (2007) demonstrate, EGM gamblers tend to travel different distances depending on the type of venue they wish to visit. In consequence, the allocation of EGM expenditure to the geographic area that contains the venue is problematic, especially for small areas containing hundreds of residents rather than thousands.

These factors make the investigation of gambling behaviour in small areas difficult. Yet understanding the spatial distribution of problem gambling is necessary for understanding and responding to inequalities in the burden of gambling-related harm. From a regulatory perspective, if the inequitable distribution of EGMs leads to a spatial concentration of gambling-related harm in poorer areas, then spatially-targeted regulations (e.g. local rather than jurisdiction-wide limits on EGM density or targeted interventions such as the removal of cash withdrawal facilities in particular areas) may be required to minimise harm. Furthermore, the spatial distribution of EGM venues should be considered a social justice issue if residents of low-income neighbourhoods receive a greater ‘exposure’ to harmful gambling products and therefore experience a greater incidence of gambling-related harms (Jerrett et al., 2009; Rosenberg, 2014). The development of improved methods for estimating the spatial distribution of gambling-related harm is especially important given the contested nature of this issue in the field of gambling studies, with critics asserting that the spatial relationship between socio-economic disadvantage and gambling found in administrative data is not supported by the findings of surveys (Delfabbro and King, 2017). There are also regulatory imperatives which make it important to understand where the impacts of gambling-related harm are most heavily experienced. For example, licensing authorities in every Australian state and territory must undertake local social impact assessments when licensing new gambling venues (see, for example, Francis et al., 2017). Yet these assessments typically take place without any local data on gambling behaviour beyond that found in administrative records. In these social impact assessments, gambling expenditure is generally not currently accepted as a proxy for gambling-related harm. The measures of harm that do exist (e.g. state-level prevalence studies) include very little geographic specificity, rendering their findings largely inconsequential in the highly localised social impact assessment process. Finally, resources for prevention and treatment services would be best targeted on the basis of local needs. Yet little is known about the spatial location of gambling-treatment needs in Australia. In short, there are imperatives from the regulatory, academic, licensing and treatment domains that make understanding the spatial distribution of problem gambling important.

Consequently, several studies to date have sought to map the spatial distribution of problem gambling risk at local spatial scales. The first study of this type investigated Video Lottery Terminals (VLTs) in Montreal. Mapping the demographic risk factors for problem gambling (i.e. gender, age, income, marital status, income, ethnicity and employment status), the authors identified a spatial correlation between areas of high risk and high levels of VLT accessibility (Robitaille and Herjean, 2008). Doran and Young (2010) took a different approach to studying problem-gambling risk in Darwin, Australia. Based on the assumption that both neighbourhood socioeconomic disadvantage and EGM accessibility are risk factors for problem gambling, they used index modelling and a Huff model to predict the location of areas of high ‘vulnerability’ for problem gambling, an approach that was replicated in Philadelphia (Conway, 2015). This same approach was modified in a study of Melbourne in which the attractiveness of EGM venues was indexed by gambling losses rather than number of EGMs (Rintoul et al., 2013). Finally, Wardle *et al.* (2016) extended the index-modelling approach using a weighted linear combination of a wide range of indicators to estimate the spatial location of problem gambling. In this study, indicators included demographic risk factors and indicators of outcomes such as the utilisation of problem gambling treatment services. This analysis resulted in a series of maps of problem gambling risk in small geographic areas in London and Manchester.

While these studies have developed a range of plausible methods for estimating spatial variation in problem-gambling risk, they are entirely predictive. None of these studies have been validated or calibrated against empirical data on gambling outcomes. Consequently, the weights that are assigned to the various elements of vulnerability indices are necessarily arbitrary. At best, the maps produced using this approach provide an educated guess regarding the location and relative prevalence of problem gambling. In consequence, there is a need for studies that can produce small area estimates of gambling-related harm that are statistically calibrated against systematically collected data on gambling outcomes.

1.1.4 Total consumption theory and gambling

One of the promising avenues for linking gambling consumption to gambling-related harm lies in the application of Total Consumption Theory to gambling (Rose & Day, 1990). However, Total Consumption Theory has been underdeveloped in the field of gambling research to date. Total Consumption Theory, or Single Distribution Theory as it is sometimes known, was first developed in the context of the study of alcohol, as drinking began to be viewed as a public health issue rather than a moral problem or a medical problem after World War II.

The dominant view of alcohol problems immediately after World War II, propounded most prominently by Jellinek (e.g. 1952), was the so-called 'disease model' of alcoholism. Under this model, 'alcohol addicts' were distinguished from those who were understood to be merely heavy drinkers. 'Addicts' were considered to be suffering a disease, characterised by loss of control over drinking, that resulted from what Jillinek called 'a predisposing *X* factor' (1952, p. 674), while heavy drinking was thought to be a class of categorically different behaviour. Put simply, in the disease model, alcoholism was a disease which afflicted a predisposed few, and was to be clearly quarantined conceptually from the non-addictive consumption of large amounts of alcohol. The cause of alcohol problems was to be found, therefore, in the predisposing factors within a small aberrant proportion of the population, rather than in social or regulatory conditions.

Total Consumption Theory of alcohol arose in response to this disease model of alcoholism. The key insight of the Total Consumption Theory of alcohol was that, as Rose and Day (1990) later put it, levels and patterns of drinking in the population as a whole affected the rates of drinking-related problems in that population. Rather than drinking-related problems deriving from predisposing factors in the individual, Total Consumption Theory challenged the disease model in two specific ways. First, it pointed out that the difference between the heaviest drinkers and those who drank less was a quantitative one rather than a qualitative one. In other words, rather than a 'predisposing X factor' distinguishing those with problems from the rest of the population, alcohol problems were viewed as a function of alcohol consumption, with those with lower degrees of consumption still at risk, albeit a lower level of risk. Second, it linked the experiences of individuals to society more broadly. It pointed out the number of high-risk drinkers was associated with the total amount of consumption in a society. 'Alcoholics' did not arise independently from the rest of the population, but were produced from them (Room and Livingston, 2017). As Norweigan scholar Ole-Jørgen Skog (1973) summarised pithily: less alcohol, fewer alcoholics.

While the hypothesis underlying Total Consumption Theory are widely credited to Ledermann in his study of alcohol consumption in France (1956; Skog, 2006), it was not until the 1970s that it gained wider currency in the field of alcohol studies in the English-speaking world. Ledermann posited that the distribution of alcohol consumption within a population would identical to the distribution within any other population with the same mean level of per capita alcohol consumption. Skog, in a series of publications, did much to bring the concept to a wider audience, providing more substantiating empirical data from national surveys and nuancing the theoretical approach (Skog, 1973). In part on this basis, Total Consumption Theory was featured heavily in 'the Purple Book' on alcohol regulation by the World Health Organisation (Bruun et al., 1975), which did much to spread the theory globally. In particular, in the Purple Book's account of Total Consumption Theory, it was pointed out that as mean consumption increased, heavy consumption concomitantly increased by an amount proportionate to the square-root of the increased mean consumption.

As the Total Consumption Theory gained currency over the subsequent decades, it had profound implications for alcohol policy and regulation. In particular, it pointed to the need for alcohol-control measures that targeted the entire population rather than focussing on high-use individuals specifically (Babor et al., 2010; Bruun et al., 1975). In this, the authors of the Purple Book were influenced by sociological research on the stigma and economic costs associated with targeting already disadvantaged individuals, rather than changing social conditions more generally (Room and Livingston, 2017). This made the theory contentious for political reasons, as it set alcohol policy researchers against the alcohol industry. While the theory suggests that alcohol-related harm is best controlled by limiting total consumption, the alcohol industry has a clear economic motive to maximise alcohol consumption. It should be no surprise, therefore, that a revived 'disease model', rebadged in terms of 'responsible drinking' is favoured by industry proponents (Babor et al., 2010).

Since the 1970s, the key tenants of the Total Consumption Theory have been empirically tested and modified. For example, it appears that the distribution of consumption is not lognormally distributed as Ledermann proposed, but is better represented by a gamma distribution (Rehm et al., 2010) – a finding that while useful for statistical modelling, does little to alter the substantive theoretical content. The most important caveat to the Total Consumption Theory that has emerged from four decades of empirical research relates to population subgroups, and the definition of a society. Specifically, it is possible for consumption among distinctive subgroups in a population to diverge from the total population, as has been occurring, for example, with regards to alcohol consumption among youth worldwide in recent years (Pennay et al., 2018). Indeed, the possibility that trends for specific subgroups may diverge from general-population trends has been evident for several decades (e.g. Herd, 1985). As such, the generality of Total Consumption Theory to all subgroups within a population should not be assumed.

Nevertheless, the empirical record has lent a great deal of support to the Total Consumption Theory of alcohol. One recent review concluded that it is time to 'accept that arguments about the distribution of alcohol consumption are largely settled in terms of the research literature—that the distribution among consumers is highly skewed and roughly lognormal, and that in this, it follows a common pattern among consumer products and behaviours' (Room and Livingston, 2017, p. 18). The main empirical postulates of the theory are consistent with the overwhelming body of evidence.

Despite its prominence in alcohol research, Total Consumption Theory has had little purchase in understanding gambling. Yet the parallels between gambling and alcohol are striking. Both are addictive forms of consumption, yet ones in which harm is also experienced by non-addicted individuals. Gambling research has its own parallel of the 'disease model of alcohol', insofar as the social impacts of gambling are usually associated only with predisposed 'problem gamblers' rather than being associated with society-wide factors. And both fields of research share similar disciplinary backgrounds, being located at the intersection of the health and social sciences. Total Consumption Theory – if applicable to gambling – also provides some particular affordances to the study of local patterns of gambling-related harm. Specifically, if the rate of gambling consumption can be directly related to the rate of gambling-related harm, then measures of consumption which are available for local areas may be used as proxies for gambling-related harm in local areas. Yet to date only a handful of studies have examined gambling using the tools provided by Total Consumption Theory (Grun and McKeigue, 2000; Hansen and Rossow, 2008; Lund, 2008). To that end, further study of Total Consumption Theory as it relates to gambling is warranted.

1.2 Research approach and aims

This thesis set out to investigate the impacts of EGMs, an entrenched social and public health issue, using health geographic approaches. In particular, it set out to develop small area estimates of problem gambling in order to inform regulation and research. It does so through a series of six, interlinked studies. Although these six papers have been written for separate publication, they collectively build an evidence base from which small area estimates of gambling-related harm can be made.

The first study aims to build a statistical model of the spatial behaviour of EGM gamblers using Huff models, calibrated against a large geocoded survey. This paper uses survey data on visitation behaviour to build a model describing the probability of people visiting specific gambling venues on the basis of their residential location.

The second, third and fourth studies aim to clarify understanding of the relationship between gambling losses and problem gambling. These studies aim to test the applicability of Rose and Day's (1990) Total Consumption Theory to gambling losses and problem gambling. Specifically, these studies investigate the relationships between problem gambling and EGM losses at the spatial scales of (a) the individual, (b) the gambling venue, and (c) the state or territory jurisdiction.

The fifth study produces small area models of problem gambling prevalence using spatial microsimulation methods. It takes a three-step approach to estimation. First, it uses the models of spatial behaviour (derived from study 1) to allocate gambling expenditure from point-geocoded EGM venues to residential small areas (median population 385). Second, it draws on the findings of the studies of the EGM loss – problem gambling relationship (studies 2, 3 & 4), to estimate the number of problem gamblers in each small area on the basis of their gambling losses. Finally, spatial microsimulation methods are used to allocate survey respondents to small areas on the basis of modelled expenditure and census-derived demographic constraints.

Having achieved the primary goal of this thesis, the sixth study aims to chart a future direction for the spatial modelling of gambling-related harm. Where the earlier studies attempted to model the prevalence of problem gambling in small areas as the outcome of interest, this study exploits spatio-temporal variation in EGM density to estimate the relationship between EGM density and an under-investigated gambling-related harm, domestic violence.

1.3 Contribution of research

The body of research presented in this thesis has advanced the state of knowledge in three specific ways. First, it has produced the first set of calibrated small area estimates of problem gambling prevalence. In order to do so, it has produced the first calibrated spatial interaction model of EGM venue visitation. It has also presented an improvement on spatial microsimulation methods, demonstrating how estimates can be improved by incorporating constraints based on administrative data sources.

Second, this thesis has contributed to the understanding of the dose-response relationship between EGM losses and problem gambling risk. Chapter 3 contests the conventional wisdom that the dose-response relationship between EGM losses and problem gambling for individuals is J-shaped, providing evidence of a linear relationship. Subsequent work demonstrates that this relationship continues to exist at the spatial scales of EGM venues and jurisdictions.

Third, this thesis takes the study of spatial relationships between EGMs and gambling-related harm beyond the study of problem gambling. It does so using the example of domestic violence, producing the first study of the spatial association between police-recorded domestic violence incidents and EGM accessibility.

1.4 Publications

The completion of a thesis 'by publication' is encouraged at The Australian National University.[2] This thesis is based on six academic journal articles, which are presented as Chapters 2 – 7 of this thesis. At the time of submission, four of these articles have been published, with the remaining two accepted for publication. All of these journal articles were co-authored, with Francis Markham the lead author on each. The respective contributions of each co-author is detailed in Appendix A.

Consequently, this thesis departs from the standard thesis format in several ways. Each article may adopt slightly different conventions. Material which is supplied in online-only supplementary appendices to the original journal article is included in this thesis in Appendices B – G. Reference lists from the original articles have been consolidated into a single reference list. When these studies refer to each other, they do so using a reference to the published article rather than a cross-reference to the chapter of this thesis. Most importantly, each article has been written to stand alone, which at times necessitates an unfortunate degree of repetition.

The journal articles that constitute Chapters 2 – 7 of this thesis are listed below.

Chapter 2

Markham, F., Doran, B. & Young, M., 2014. Estimating gambling venue catchments for impact assessment using a calibrated gravity model. *International Journal of Geographical Information Science*, 28(2), p.326–342.

Chapter 3

Markham, F., Young, M. & Doran, B., 2016. The relationship between player losses and gambling-related harm: evidence from nationally representative cross-sectional surveys in four countries. *Addiction*, 111(2), p.320–330.

Chapter 4

Markham, F., Young, M. & Doran, B., 2014. Gambling expenditure predicts harm: evidence from a venue-level study. *Addiction*, 109(9), p.1509–1516.

[2] The ANU procedure for submitting and examining theses by compilation is available at https://policies.anu.edu.au/ppl/document/ANUP_003405

Chapter 5

Markham, F., Young, M., Doran, B., Sugden, M., 2017. A meta-regression analysis of 41 Australian problem gambling prevalence estimates and their relationship to total spending on electronic gaming machines. *BMC Public Health*, 17(495), p.1–11.

Chapter 6

Markham, F., Young, M., Doran, B., 2017. Improving spatial microsimulation estimates of health outcomes by including geographic indicators of health behaviour: The example of problem gambling. *Health & Place*, 46(2017), 29–36.

Chapter 7

Markham, F., Doran, B. & Young, M., 2016. The relationship between electronic gaming machine accessibility and police-recorded domestic violence: A spatio-temporal analysis of 654 postcodes in Victoria, Australia, 2005–2014. *Social Science & Medicine*, 162, p.106–114.

Chapter 2: Estimating gambling venue catchments for impact assessment using a calibrated gravity model

2.1 Foreword

The aim of this chapter was to test the proposition that the spatial behaviour of gamblers – that is, their choice of which EGM venues to visit – could be modelled in a Geographical Information System (GIS). Previous research had demonstrated that the spatial behaviour of gamblers was neither necessarily straightforward nor consistent (Doran et al., 2007). While some EGM venues were described as drawing their patrons from several kilometres away, others had catchments of limited spatial extent.

This variability has implications for both social impact assessment and the production of small area models of problem gambling. In the case of social impact assessment, it is difficult to know the geographical scope of social impacts when catchments are so variable in extent. In the case of the production of small area estimates, it is well documented that access to gambling venues is a risk factor for problem gambling (Pearce et al., 2008; e.g. John W Welte et al., 2004). Yet it is difficult to understand how proximity translates into risk without knowledge of the nature of spatial relationship between place of residence and EGM venue.

Consequently, this study set out to model EGM venue visitation behaviour on the basis of a geocoded postal survey of residents which asked questions about which specific EGM venues respondents visited. To do so, it used the well-understood Huff model, a type of spatial interaction model that is sometimes referred to as a gravity model. It was especially successful in modelling the venue visitation behaviour of EGM gamblers. The Huff model produced in this study has been incorporated into the estimates undertaken in Chapters 4 and 6. As such, it is integral to the studies in the remainder of the thesis.

This chapter was published as:

Markham, F., Doran, B. & Young, M., 2014. Estimating gambling venue catchments for impact assessment using a calibrated gravity model. *International Journal of Geographical Information Science*, 28(2), p.326–342.

2.2 Abstract

Gambling using electronic gaming machines has emerged as a significant public health issue. While social impact assessments are required prior to the granting of new gaming licenses in Australia, there are few established techniques for estimating the spatial distribution of a venue's clientele. To this end, we calibrated a Huff model of gambling venue catchments based on a geocoded postal survey (*n* = 7,040). We investigated the impact of different venue attractiveness measures, distance measures, distance decay functions, levels of spatial aggregation, and venue types on model fit and results. We then compared model estimates for different behavioural subgroups. Our calibrated spatial model is a significant improvement on previously published models, increasing R^2 from 0.23 to 0.64. Venue catchments differ radically in size and intensity. As different population subgroups are attracted to different venues, there is no single best index of venue attractiveness applicable to all subpopulations. The calibrated Huff model represents a useful regulatory tool for predicting the extent and composition of gambling venue catchments. It may assist in decision making with regard to new license applications and evaluating the impact of health interventions such as mandated reductions in EGM numbers. Our calibrated parameters may be used to improve model accuracy in other jurisdictions.

2.3 Background

Gambling is a significant public health issue wherever commercial gambling opportunities are widely available. Estimates of gambling-related harm in the general population of Western countries range from 1.8% in Australia (Productivity Commission, 1999) to 7.8% in Canada (Currie et al., 2006) depending on the measure used. At the level of the individual, the harms associated with gambling may include psychiatric problems, suicide, alcohol and drug problems, financial problems, and criminal behavior (Korn and Shaffer, 1999). Other gambling-related harms such as regressive distribution of economic resources are social determinants of health and wellbeing that operate at the community level (Productivity Commission, 1999).

Gambling liberalisation during the last thirty years has resulted in the proliferation of commercial gambling opportunities in many developed countries. In the United States, for example, the number of states that authorise casino gambling rose from two in 1988 to thirty-eight in 2011 (American Gaming Association, 2012; Eadington, 1998). Similarly, in Australia, the number of electronic gaming machines (EGMs, the Australian variant of the slot machine) increased from 48,439 to 198,725 in the thirty years to 2010 (Office of Economic and Statistical Research, 2012; Wilkinson, 1996).

Increased gambling accessibility has resulted in a rise in the prevalence of gambling-related harms (Shaffer et al., 1999; Storer et al., 2009). In Australia, EGM density has been closely associated with elevated rates of gambling harm (Productivity Commission, 1999). When considered at the scale of the state jurisdiction, the availability of EGMs in venues other than casinos was associated with a tripling of the prevalence of problem gambling (Productivity Commission, 1999), with each additional 100 EGMs associated with 79 new problem gamblers (Storer et al., 2009).

Given that EGMs are a venue-based form of gambling, opportunities exist to intervene at the venue level to reduce EGM-related harm. While reversing the trend of increased gambling accessibility is politically challenging, other venue-level interventions such as self-exclusion programs, limits on access to automatic teller machines, and caps on machine numbers in venues, have become routine. In particular, every jurisdiction in Australia mandates that social impact assessments be undertaken prior to the granting of a new gaming machine license in order to allow the harmful impacts of venue-based gambling to be considered.

To assess the social impact of individual venues effectively we need to know the spatial distribution of the venue's clientele – the people most directly impacted by the venue. However, surprisingly little is known about the extent and intensity of the 'catchments' of gambling venues – that is, the spatial coverage of the catchments and the proportion of residents within them who visit gambling venues. To date, only two studies have specifically investigated the geography of gambling venue catchments. KPMG Consulting (2000) asked survey respondents how far they travelled to the gambling venue on the last occasion they gambled on EGMs in Victoria, Australia. From this data, KPMG produced an averaged 2.5 km radial catchment for each venue. However, this estimate is problematic because it relies on the dubious assumptions that (a) respondents can reliably estimate their own travel distances (cf. Walmsley and Jenkins, 1992) and b) the size and shape of catchments is identical across all venues. In contrast, Doran *et al*. (2007) performed a geocoded household survey that asked respondents to identify their preferred gambling venue. Network distance was calculated from each respondent's residence to his or her preferred gambling venue. This study found considerable variation in catchment radius, from over 14 km to less than 4 km.

Both of these studies assumed that the catchment areas of venues are constant across all groups of visitors. However, there is reason to doubt this claim. Young *et al*. (2012a) found that EGM gamblers and problem gamblers are more likely to visit venues closer to their homes than non-gamblers and non-problem gamblers. If catchment sizes differ between groups of visitors, then a 'one size fits all' approach to catchment estimation may not be appropriate for social impact assessment.

While the findings produced by geocoded population surveys are able to provide catchment information for use in social impact assessments, they may be prohibitively expensive to conduct. Of greater utility would be a predictive tool that could accurately estimate venue catchments for a range of venue sizes using secondary, freely available, data. Here the gravity modelling approach developed in retail and trade geography may be useful. In particular, the Huff model (1964) has been used for over four decades to probabilistically estimate the market areas of retail outlets, and is still considered the best tool for this purpose in conjunction with contemporary Geographic Information Systems (GIS) (Huff and McCallum, 2008).

Three studies have employed the Huff model to investigate gambling venue catchments. Doran and Young (2010) developed a local measure of gambling accessibility using a Huff model at the city-scale. Their analysis was replicated for metropolitan Melbourne by Rintoul *et al.* (2013), who sought to demonstrate that gambling expenditure was associated with EGM accessibility and to identify localities where harm-minimisation efforts are most urgently needed. At the national scale, Markham *et al.* (2014a) employed the Huff model to estimate the catchment areas of casinos in Australia. None of these studies empirically calibrated model parameters, instead using parameters selected *a priori* from the trade-area modelling literature. Despite the sound theoretical basis of Huff models, this is problematic because model accuracy is highly dependent on the parameters chosen. To date, no study has calibrated a trade-area model for gambling against venue visitation data, nor have the predictions of these models been compared to actual gambler behaviour. Consequently, the utility of the Huff model in a gambling context remains untested. To address this shortfall, we calibrated a Huff model to estimate gambling venue catchments using a large, geocoded postal survey. We posed three specific research questions:

1. To what extent can Huff models predict the spatial distribution of gambling venue patrons?
2. Which parsimonious configuration of model parameters provide the best model fit with observed visitation data?
3. Does venue attractiveness vary between population subgroups?

We empirically assessed the suitability of the Huff model for explaining gambling venue visitation patterns using goodness-of-fit indices. Results of this calibrated model were compared with those from a previous normative study (Doran and Young, 2010). We then compared Huff models for visitor subgroups to investigate differences in venue-choice behaviour.

2.3.1 Study-area

As part of a larger project, this paper builds on a trajectory of gambling research in the Northern Territory (NT) of Australia (Doran and Young, 2010; e.g. Young et al., 2009). The NT is notable for its relatively small population (229,711 in 2010), geographic remoteness, and relatively high proportion of Indigenous residents (30%, compared to 3% in the rest of Australia). This study specifically focused on the three largest towns in the NT, which contain an estimated 63% of its population: Darwin (107,430 persons), Alice Springs (27,987 persons) and Katherine (10,104 persons) (Australian Bureau of Statistics, 2011). EGMs in the NT are concentrated in these three towns, which hosted 88% of the jurisdiction's EGMs (n = 1,798) in June 2010, but just 63% of its population. While 46% of the EGMs in these towns are located in casinos in Darwin and Alice Springs (833 EGMs), gambling opportunities are dispersed across the study site, with EGMs available in 26 clubs (612 EGMs) and 36 hotels (353 EGMs). Clubs, such as sporting or returned servicepersons clubs, are not-for-profit entities restricted to a cap of 45 EGMs per venue. Hotels or pubs are private businesses capped at 10 EGMs per venue.

It is important to recognise that the Northern Territory may not be entirely representative of Australia. For example, the Indigenous population of the Northern Territory is over 30%, compared to the rest of Australia where the Indigenous population is around 3%. However, the survey was designed to mitigate the geographic particularities of the Northern Territory context. The areas surveyed have a much lower proportion of the population who are Indigenous. Furthermore, Indigenous people are generally less likely to respond to mail surveys than the general population. In this way, while Indigenous non-response may make a survey less representative of the population in the Northern Territory, it is likely to make the survey more representative of Australia as a whole.

Furthermore, administrative data suggests that EGM gambling in the Northern Territory is not substantially dissimilar to Australia as a whole. As Figure 1.1 showed, per capita annual EGM expenditure in the Northern Territory was \$460 in 2015-16, compared to a national average of \$630. While this is below average, Tasmania had much a lower per capita EGM expenditure of \$280 per annum, while per capita expenditure in New South

Wales was well above average ($980 per annum). In other words, while expenditure in the Northern Territory is below average, it is not qualitatively different to the rest of the country in this regard. Furthermore, as Figure 1.2 showed, EGM participation rates – as revealed by surveys – is approximately average for Australian states and territories. Combining these two figures, per EGM gambler annual expenditure in the Northern Territory is $2000 per year, similar to the equivalent figure in Queensland. In short, EGM gambling expenditure and participation in the Northern Territory are not exceptional, and are approximately in line with national averages.

2.4 Data and Methods

We used the Huff model to estimate gambling venue catchments. The Huff model is a form of spatial interaction model, which seeks to describe in a spatially explicit manner flows of people across space to a fixed set of locations in order to access goods or services. The Huff model takes the form:

$$P_{ij} = k \cdot o_j^{\gamma} \frac{\left(\prod_l a_{il}^{\alpha_l}\right) \cdot f(d_{ij}, \beta)}{\sum_i \left[\left(\prod_l a_{il}^{\alpha_l}\right) \cdot f(d_{ij}, \beta)\right]} \tag{1}$$

where P_{ij} is the probability of residents at origin j interacting with destination i; o is the population of origin j; a_{il} is the lth variable describing the attractiveness of destination i; f is a function of the distance between origin j and destination i; and k, γ, α_l, and β are parameters to be empirically estimated. When fit using actual flow data, these estimated parameters can be interpreted in a similar manner to the output of ordinary linear regression, with the outcome variable representing the estimated probability of interaction between a source and destination, conditional on a set of predictor variable values. Like ordinary linear regression, the Huff model can be used to describe patterns in a data set, test hypotheses or make predictions.

The Huff model calibration process requires both population-level gambling venue visitation data and venue-level attractiveness data (see Figure 2.1). We used postal methods to collect venue visitation data and compiled venue-level attractiveness variables from secondary data sources. We then calculated distances between respondent residential locations and gambling venues and aggregated survey responses using two different spatial zoning schemes. We assessed a series of Huff model configurations using different distance measures, combinations of attractiveness variables, and distance decay functions to determine which of a variety of parameter configurations to include in our catchment model. Finally, using the model configuration identified as most appropriate, we estimated venue catchments for a series of patron subgroups likely to be of interest to gambling regulators.

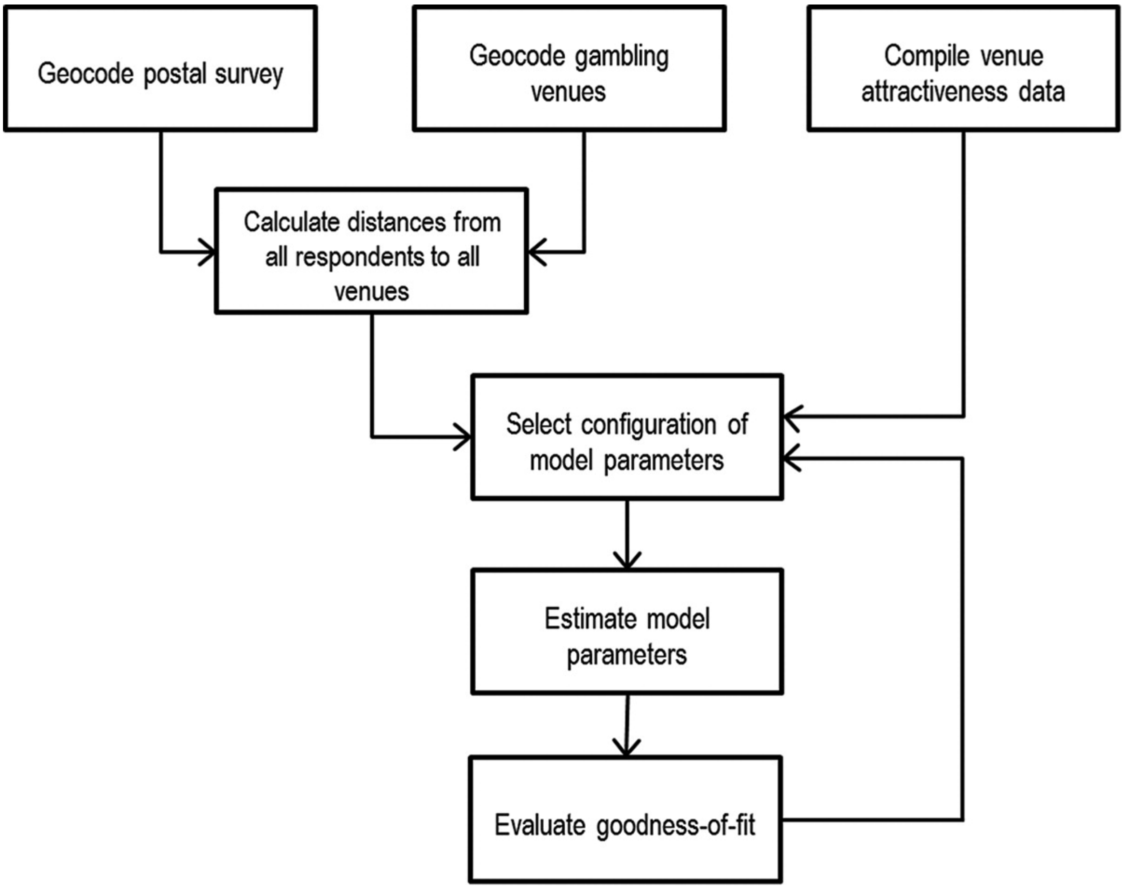

Figure 2.1: Huff model calibration process diagram

2.4.1 Visitation data

Using the geocoded national address file (or G-NAF: PSMA Australia, 2010) as a sample frame, we conducted a postal survey of 46,263 addresses in Darwin, Katherine and Alice Springs. The G-NAF is an authoritative database of verified geocoded street addresses for Australia, collated from various government agency databases including those of the Australian Electoral Commission and Australia Post. We mailed questionnaires to all G-NAF addresses to which Australia Post would deliver unsolicited mail and which were zoned as residential. To extend our spatial coverage, we selected 2,300 addresses across the peri-urban fringes of Alice Springs and Darwin, to which Australia Post does not deliver mail, for hand delivery of questionnaires. The authors drove to the selected addresses and pegged questionnaires with reply-paid envelopes to gates and fences. In Alice Springs, 300 hand-deliveries were conducted within a 15 km radius of the CBD. In the Darwin peri-urban fringe, we used a spatially stratified cluster sample design to select 2,000 out of a potential 7,000 addresses in a band 20 km to 40 km from Darwin's Central Business District. For the purpose of sampling, we divided this spatial band into four concentric tracts, each 5 km wide and selected 500 addresses for hand delivery in each tract, grouped into several contiguous blocks. The questionnaires were mailed between April and August 2010 and hand delivered to Alice Springs and Darwin in July and September 2010, respectively. Any household member aged eighteen or older was eligible to respond, and return of the survey implied consent. The Human Research Ethics Committee of Charles Darwin University granted approval to conduct the study (protocol no. H09048).

The questionnaire asked which gambling venues the respondent had visited in the last month. Respondents selected their most frequently visited venue from a list of all EGM venues in or proximate to their town of residence. Participants were asked to report the number of times they had visited this venue in the last month, and whether they participated in EGM gambling on their last visit. A unique identifier that referenced the respondent's G-NAF record was also included on the questionnaire, enabling survey returns to be precisely geocoded.

Participants were asked to complete the Problem Gambling Severity Index (PGSI) for the last twelve-months (Ferris and Wynne, 2001a). We used the PGSI as our measure of gambling-related harm as it is a clinically-validated scale used to estimate problem-gambling risk in the general population (Ferris and Wynne, 2001a; Neal et al., 2005). It asks respondents to answer a series of nine questions. Possible answers to these questions and there attendant scores are:

- never (score: 0)
- rarely (score: 1)
- often (score: 2)
- always (score: 3)

Respondents were asked to answer the following questions with respect to their gambling in the last twelve months:

1. Have you bet more than you could really afford to lose?

2. Have you needed to gamble with larger amounts of money to get the same feeling of excitement?
3. Have you gone back on another day to try to win back the money you lost?
4. Have you borrowed money or sold anything to gamble?
5. Have you felt that you might have a problem with gambling?
6. Have people criticised your betting or told you that you had a gambling problem, whether or not you thought it was true?
7. Have you felt guilty about the way you gamble or what happens when you gamble?
8. Has gambling caused you any health problems, including stress or anxiety?
9. Has your gambling caused any financial problems for you or your household?

The answers to these questions are summed to produce a PGSI score. Responses range on an ordinal scale from 0-27, routinely classified into groups having no risk (PGSI 0), low risk (PGSI 1-2), moderate risk (PGSI 3-7) or high risk (PGSI 8+) of being a problem-gambler.

2.4.2 Venue attractiveness data

For each venue, we obtained data on type of gaming license (i.e. hotel, club or casino), number of EGMs licensed to the venue on June 30 2010, and street addresses from the Northern Territory Department of Justice. We manually geocoded venue addresses using Google Maps. We also selected several venue-level spatial variables relevant to venue visitation behaviour: proximity to centres of community congregation, distance from the Central Business District (CBD) as measured by distance to the general post office, participation in the tourism-oriented night-time economy, and proximity to the ocean. In particular, proximity to centres of community congregation has been suggested as an important predictor of gambling catchments as these venues have potential to tap the pre-existing activity spaces of large numbers of residents (e.g. Doran et al., 2007). We used road network distance to closest supermarket as a proxy measure of proximity to areas of community congregation and obtained supermarket location data from the websites of the two supermarket operators who collectively supply 75% of the grocery market (Australian Competition and Consumer Commission, 2008). Supermarket data were collected in 2011. We included distance from the CBD in this study due to the differences in EGM gambling markets between suburban venues and inner city venues (e.g. Young et al., 2009). We measured distance to the CBD by proxy as road network distance to the general post office. Proximity to the ocean was included in this model due to our observation during exploratory data analysis that venues located within 100 m of the ocean had more patrons than might be otherwise expected, an observation consistent with previous research conducted at a coarser spatial scale (Wardle et al., 2014). We defined venues as proximate to the ocean if they were located within 100 m of the coastal boundary. This measure was adopted on the basis of exploratory modelling. We followed this up with venue visits, and noted that a large number of patrons to these venues were visiting primarily to access the amenity of ocean views, rather than to gamble. On this basis, venue visitorship was higher than would be expected on the basis of gambling products alone. Participation in the night-time tourist economy was measured on the basis of observation and local knowledge.

While Rintoul *et al.* (2013) used gambling expenditure as the measure of venue attractiveness in their Huff model, we considered this inappropriate for a study concerned with estimating general visitation. Specifically, because visitation is a necessary condition for EGM gambling, explanation of visitation behaviour based on gambling expenditure is temporally inconsistent and violates model assumptions.

2.4.3 Distance calculations, spatial aggregation and parameter estimation

The Huff model can be extended to incorporate different measures of distance. We calculated the distance between each survey respondent and EGM venue using Euclidian, Manhattan and road network measures of distance (in kilometres). We excluded respondents who did not report visiting any venue in the past month. We aggregated individual responses into Mesh Blocks and census collector districts in order to test which level of spatial aggregation would provide the best model fit without biasing parameter estimates. The distance from each zone to each venue for the three distance measures was estimated by calculating distances at the household level and then taking the median in each zone (Batty and Sikdar 1982).

Mesh blocks are a micro-level geographical unit, with a size of 20-50 dwellings in residential areas. Mesh blocks were chosen as the origin zones in this study as their relatively small size minimizes the effect of the modifiable areal unit problem (MAUP). Census collector districts correspond to the area assigned on census night to a single census officer, and contain an average of seven mesh blocks in our study area. In our case, the MAUP – that is, bias resulting from the use of arbitrary administrative zones to aggregate respondents – might make it difficult to know whether parameter estimates are the result of actual travel behavior or just the choice of zoning system (Openshaw, 1977). Although more susceptible to the effects of the MAUP, collector districts were included in this study in order to increase the number of survey respondents per unit. When comparing systems of spatial aggregation at different scales, increasing average zone size will generally increase measures of model performance (Batty and Sikdar, 1984). However, this does not necessarily mean that a better estimate of visitation behaviour has been derived. Rather, because larger zones will contain more responses, aggregation ameliorates the ‘small numbers problem’ whereby visitation patterns in zones with few respondents appears to be increasingly random (Batty and Sikdar, 1982). Consequently, we compared the parameter estimates at different levels of spatial aggregation to see if the MAUP resulted in model bias.

We employed a composite measure of venue attractiveness following the observation in the retail trade-area literature that composite measures are more accurate predictors of shopping behaviour than centre size alone (Gautschi, 1981). Our measure included license type, number of EGMs, logarithm-transformed distance to supermarket, logarithm-transformed distance to CBD, having ocean views, and being a tourist-oriented inner-city bar. License type, ocean views and inner-city bar variables were coded as integer variables taking the values of one for false and two for true. We calibrated our model against a matrix of respondents’ most frequently visited venue, coded one if the venue was the preferred venue or zero if not.

Parameters were estimated using maximum likelihood methods using the *R* software package (R Development Core Team, 2012). We maximised the log-likelihood equation derived by Fotheringham and O’Kelly (1989) and computed confidence intervals for estimated parameters from the covariance matrix obtained by inverting the optimised hessian matrix. Goodness-of-fit was calculated using the R^2 and Standardised Root Mean Square Error (SRMSE) metrics suggested by Fotheringham and O’Kelly (1989) and Thorsen and Gitlesen’s (1998) Relative Number of Wrong Predictions (RNWP).

2.4.4 Model selection

In order to investigate which configuration of variables, distance measures, distance decay functions and venue-type subsets are most useful for predicting EGM gambling catchments we tested ten different models (see Table 2.1). Specifically, Model 1 was set to match the parameters used in the study published by Doran and Young (2010). Model

2 used the same configuration of variables, but allowed parameters to be estimated from the survey data. Subsequent models 3–9 were based on this configuration but each varied a single configuration option. Model 3 tested the effect of the removal of attractiveness index entirely. Model 4 introduced a composite attractiveness index. Model 5 modified Model 2 by aggregating responses to census collector districts. Models 6 and 7 tested the effect of using network and Manhattan measures of distance, respectively. Model 8 used an exponential distance decay function in place of the power function. Model 9 included casinos and their visitors in the specification. Finally, in Model 10, we combined the best-fitting variants of the previously tested models.

2.4.5 Comparison of visitor subgroups

We re-estimated the best fitting Huff model configuration (Model 10) using visitation data for different subgroups of venue visitors. The subgroups of interest included:

1. EGM gamblers (respondents who participated in EGM gambling on their last visit to their most frequently visited venue)
2. non-gamblers (respondents who did not gamble on their last visit)
3. moderate- to high-risk visitors (respondents with a PGSI score of three or more)
4. non-problem gamblers (respondents with a PGSI score of zero)
5. walkers (respondents who travelled on foot)
6. frequent visitors (respondents who visited their most frequently visited venue four or more times in the last month)
7. infrequent visitors (respondents who visited their most frequently visited venue once in the last month)

We selected these subgroups based on their relevance to regulators and our untested hypothesis that they would illustrate divergent travel behaviours. A low-income subgroup was not separately modelled because of the large amount of missing data on this variable. We mapped the estimated catchments for visitors from these different subgroups using kernel density estimation with a bandwidth of 500m, using ArcGIS 9.3 (ESRI, 2010). The 500m bandwidth was selected on the basis of an exploratory analysis, aimed at producing a resolution of patterns that was of approximately the same spatial resolution as the underlying meshblock geometry. We selected three well-known venues in central and northern Darwin to map based on their diversity in terms of locational, licensing and patron characteristics. These three venues were chosen from all the venues in the study area (*n* = 64) for illustrative rather than analytic purposes and are indicative of the variation among venues.

2.5 Results

2.5.1 Survey respondents

We received 7,040 survey responses (14.5% response rate), with a median of 6 responses per Mesh Block (*IQR* = 4–8) or 32 responses per census collector district (*IQR* = 23–44). Because all addresses in the sample frame were already geocoded, we achieved a 100% geocoding match rate. There was little evidence of spatial clustering of response rates at the Mesh Block level (Moran's $I = 0.003$). More detail of the sampling strategy and response rate for different geographical areas is given in Appendix B, with Katherine recording a lower response rate for reasons that are unclear.

As is typical for surveys of this kind, the sample was older (mean age = 48.9) and contained a higher proportion of women (61.8%, $n = 4{,}292$) compared to the population of the same area in the 2006 Census of Population and Housing (median age = 30-44, proportion of women = 48.5%). The majority of respondents were residents of Darwin (77.3%, $n = 5442$), with the remaining respondents residing in Alice Springs (19.8%, n =1393) and Katherine (2.9%, n =205). Among the 71.1% ($n = 4{,}857$) of respondents who had visited an EGM venue in the last month, 20.9% ($n = 1{,}013$) gambled on EGMs during their last visit. In terms of gambling-related harm 4.6% ($n = 324$) were at moderate risk (PGSI 3-7) and a further 2.0% ($n = 143$) were at high risk (PGSI 8+) of problem gambling. While this is a substantially higher level of gambling-related harm than that found in the most recent NT prevalence survey (0.64% PGSI 8+, Young et al., 2006), this may be accounted for by incommensurate survey methods and relatively large standard errors. A more detailed summary of survey responses is provided in Table 2.2.

2.5.2 Model selection

The Huff model was able to explain aggregate community-venue visitation patterns to a moderate degree (see Table 2.1). The best fitting model (Model 10, $R^2 = 0.64$) aggregated responses to the census collection district level, used a Euclidian distance measure, a power distance decay function, included casino venues, and included a variety of situational and licensing attractiveness variables. This represents a substantial improvement in explanatory power when compared to the use of Doran and Young's (2010) parameters (Model 1, $R^2 = 0.23$). Even when retaining Doran and Young's model configuration but empirically estimating parameters a substantial improvement in model fit was achieved (Model 2, $R^2 = 0.31$).

In terms of individual attractiveness indices, the number of EGMs at a venue was a useful predictor of venue attractiveness. Venues with 45 EGMs were estimated to be 6.5 times as attractive as venues with a single EGM when number of EGMs was used as the only attractiveness variable (Model 2). The introduction of other licensing and situational variables increased the estimated magnitude of this relationship, with venues with 45 EGMs estimated to be 14.9 times as attractive as venues with a single EGM, holding other attractiveness variables constant (Model 4). As the EGM α is estimated to be less than 1.0 in all cases, the attractiveness of a venue does not increase linearly with addition of new EGMs as assumed by previous studies (Model 1).

While the model provides an adequate fit to the data when aggregated at the Mesh Block level (Model 2, $R^2 = 0.31$), CD level aggregation dramatically improves model performance (Model 5, $R^2 = 0.55$). Parameter estimates did not vary significantly when aggregating (see Models 2 and 5), indicating that in this case aggregation is unlikely to induce MAUP-related bias.

The use of network distance or Manhattan distance metrics had little impact on model fit (Models 6 & 7). A distance decay power function better fitted the interaction data than an exponential function ($R^2 = 0.28$, Model 8), an unusual result for intra-urban interactions (Fotheringham and O'Kelly, 1989).

The inclusion of casinos in the model improved overall model fit ($R^2 = 0.36$, Model 9). Respondents visited casinos for reasons over and above the number of EGMs they house, with Model 9 estimating casinos to be 1.7 times as attractive as a hypothetical non-casino venue with the same number of EGMs. Investigation of Variance Inflation Factors (VIFs) among the attractiveness variables revealed substantial levels of covariance between the number of EGMs in a venue and the venue's status as a casino (but not other pairs of attractiveness variables). The collinearity between these variables results in changing parameter estimates between models and means that these estimates should be treated with caution.

2.5.3 Comparison of visitor subgroups

Estimated gambling venue catchments varied substantially across subgroups (see Table 2.3). Compared to all visitors (Group A, $\beta = 1.02$), distance decay was higher among those who gambled on EGMs on their last visit to a venue (Group B, $\beta = 1.18$), those at moderate or greater risk of problem gambling (Group D, $\beta = 1.16$), those who walked to a venue (Group F, $\beta = 1.\ 71$), and frequent visitors to a venue (Group G, $\beta = 1.13$).

Table 2.1: Huff model parameter estimates for different configurations of parameters

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Include casinos	No	No	No	No	No	No	No	No	Yes	Yes
Zone	Mesh blocks	Mesh blocks	Mesh blocks	Mesh blocks	CCDs	Mesh blocks	Mesh blocks	Mesh blocks	Mesh blocks	CCDs
Distance type	Euclidian	Euclidean	Euclidean	Euclidean	Euclidean	Network	Manhattan	Euclidean	Euclidean	Euclidean
Decay function	Power	Power	Power	Power	Power	Power	Power	Exponential	Power	Power
β	1.5	**1.09 (1.06, 1.12)**	**1.14 (1.11, 1.17)**	**1.08 (1.05, 1.11)**	**1.13 (1.09, 1.16)**	**1.17 (1.13, 1.20)**	**1.07 (1.04, 1.10)**	**0.17 (0.17, 0.18)**	**1.02 (0.99, 1.05)**	**1.02 (0.99, 1.05)**
EGMs α	1	**0.49 (0.44, 0.54)**		**0.68 (0.61, 0.76)**	**0.48 (0.43, 0.53)**	**0.47 (0.42, 0.52)**	**0.50 (0.45, 0.55)**	**0.56 (0.51, 0.61)**	**0.52 (0.48, 0.57)**	**0.71 (0.64, 0.77)**
Casino α									0.73 (0.47, 0.99)	-1.72 (-2.14, -1.29)
Club α				-0.00 (-0.15, 0.14)						-0.33 (-0.46, -0.19)
ln dist. to supermarket α				-0.61 (-0.84, -0.39)						-0.35 (-0.56, -0.14)
ln dist. to CBD α				**1.16 (0.81, 1.51)**						0.70 (0.41, 1.00)
Proximate to ocean α				**2.65 (2.49, 2.81)**						1.91 (1.76, 2.06)
Inner city bar α				**0.89 (0.63, 1.15)**						0.52 (0.29, 0.75)
γ	1	1.00 (0.95, 1.05)	1.00 (0.95, 1.05)	1.00 (0.94, 1.05)	1.00 (0.94, 1.06)	1.00 (0.95, 1.05)	1.00 (0.95, 1.05)	1.00 (0.95, 1.05)	1.00 (0.95, 1.05)	0.94 (0.89, 1.00)
k	1	**0.98 (0.97, 0.99)**	**0.96 (0.95, 0.98)**	**0.97 (0.96, 0.98)**	**0.99 (0.98, 1.01)**	0.99 (0.98, 1.00)	0.99 (0.98, 1.00)	1.12 (1.11, 1.13)	**0.91 (0.90, 0.92)**	**1.16 (1.15, 1.18)**
R^2	0.23	0.31	0.27	0.35	0.55	0.31	0.31	0.28	0.36	0.64
SRMSE	4.8	4.5	4.6	4.4	2.5	4.5	4.5	4.6	4.1	2.2
RNWP	1.3	1.4	1.4	1.3	1.0	1.4	1.4	1.5	1.3	0.8
n respondents	3907	3907	3907	3907	3907	3907	3907	3907	4977	4977

Note: Cells shaded grey indicate changes to model selection compared to Model 1, which was derived from Doran and Young (Doran and Young, 2010). 95% confidence intervals for parameter estimates are indicated in parentheses. Bold cells indicate 95% confidence intervals that do not overlap zero for β and α or one for γ and *k*. No confidence intervals are indicated for Model 1 because these parameters were input based on previous research, instead of being estimated from the data. An odds ratio is obtained for a binary attractiveness variable α using the formula 2^{α}. The attractiveness contribution of a continuous attractiveness variable α with value x is obtained by the expression x^{α}. CCDs refer to census collector districts.

Table 2.2: Summary of survey responses

Variable	Value
Mean age (SD)	49 (14)
Women (%)	4292 (61.8)
Mean number of times visited a venue (SD)	2.4 (3.8)
Played pokies in last month (%)	1013 (14.4)
PGSI non-problem (%)	6055 (86.0)
PGSI low risk (%)	518 (7.4)
PGSI moderate risk (%)	324 (4.6)
PGSI high risk (%)	143 (2.0)
Resident of Alice Springs (%)	1393 (19.8)
Resident of Darwin (%)	5442 (77.3)
Resident of Katherine (%)	205 (2.9)
Drove own vehicle to venue (%)	3832 (54.4)
Walked to venue (%)	507 (7.2)
Took a lift with someone else to venue (%)	407 (5.8)
Rode a bicycle to venue (%)	116 (1.6)
Caught a taxi to venue (%)	166 (2.4)
Took a public bus to venue (%)	110 (1.6)
Took a courtesy bus to venue (%)	10 (0.1)

Attractiveness parameter estimates were similarly variable among subgroups. Compared to all visitors (Group A, $\alpha = 0.71$), the number of EGMs was a more important predictor of attractiveness for those who gambled on EGMs on their last visit to a venue (Group B, $\alpha = 1.17$) and a less important predictor of attractiveness for non-gamblers (Group C, $\alpha = 0.41$), non-problem gamblers (Group E, $\alpha = 0.47$), those who walked to a venue (Group F, $\alpha = 0.37$), and frequent venue visitors (Group G, $\alpha = 0.40$). Other attractiveness variable estimates also fluctuated between subgroups. For example, while proximity to the ocean was not a significant predictor of attractiveness for EGM gamblers, non-gamblers were 5.1 times more likely to visit venues with ocean views than those without (Group C, $\alpha = 2.36$).

The venue catchment maps revealed stark differences in catchment size and intensity between venues. Among all venues in Darwin, the SKYCITY Casino (Figure 2.2) had the largest and most intense catchment for every subgroup of visitors, with the exception of those travelling on foot. Indeed, for all visitors (Group A) SKYCITY

Table 2.3: Huff model parameter estimates for different subgroups of visitors

	Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H
Subgroup	All visitors	EGM gamblers	Non-gamblers	PGSI ≥ 3	PGSI = 0	Walkers	Frequent visitors	Infrequent visitors
β	**1.02 (0.99, 1.05)**	**1.18 (1.09, 1.27)**	**1.01 (0.97, 1.05)**	**1.16 (1.04, 1.28)**	**1.06 (1.03, 1.10)**	**1.71 (1.58, 1.84)**	**1.13 (1.07, 1.19)**	**0.93 (0.88, 0.99)**
EGMs α	**0.71 (0.64, 0.77)**	**1.17 (0.98, 1.36)**	**0.41 (0.34, 0.48)**	**0.84 (0.61, 1.08)**	**0.47 (0.40, 0.53)**	**0.37 (0.11, 0.63)**	**0.40 (0.29, 0.51)**	**0.45 (0.34, 0.56)**
Casino α	**-1.72 (-2.14, -1.29)**	-0.23 (-1.37, 0.92)	**-0.95 (-1.46, -0.44)**	-0.59 (-2.11, 0.93)	**-0.59 (-1.04, -0.14)**	-1.82 (-3.66, 0.03)	0.05 (-0.69, 0.79)	-0.31 (-1.03, 0.41)
Club α	**-0.33 (-0.46, -0.19)**	0.12 (-0.36, 0.61)	0.01 (-0.14, 0.16)	-0.15 (-0.71, 0.41)	-0.09 (-0.23, 0.04)	-0.26 (-0.75, 0.23)	**0.25 (0.01, 0.49)**	-0.21 (-0.43, 0.01)
ln dist. to supermarket α	**-0.35 (-0.56, -0.14)**	-0.31 (-0.96, 0.34)	**-0.41 (-0.67, -0.14)**	-0.52 (-1.32, 0.29)	-0.00 (-0.24, 0.23)	**-1.59 (-2.30, -0.87)**	**-0.48 (-0.87, -0.08)**	-0.37 (-0.73, -0.00)
ln dist. to CBD α	**0.70 (0.41, 1.00)**	-0.26 (-1.07, 0.56)	**-0.47 (-0.79, -0.15)**	-0.26 (-1.31, 0.80)	**-0.50 (-0.79, -0.20)**	**-1.19 (-2.11, -0.27)**	-0.17 (-0.69, 0.34)	**-0.58 (-1.05, -0.10)**
Proximate to ocean α	**1.91 (1.76, 2.06)**	0.20 (-0.33, 0.74)	**2.36 (2.19, 2.53)**	**1.21 (0.59, 1.83)**	**1.99 (1.84, 2.15)**	**2.20 (1.64, 2.77)**	**1.56 (1.28, 1.84)**	**2.16 (1.92, 2.40)**
Inner city bar α	**0.52 (0.29, 0.75)**	-0.18 (-0.95, 0.59)	0.16 (-0.10, 0.41)	-0.22 (-1.13, 0.69)	**-0.40 (-0.65, -0.15)**	**1.08 (0.48, 1.69)**	-0.26 (-0.68, 0.16)	**-1.15 (-1.59, -0.71)**
γ	0.94 (0.89, 1.00)	1.01 (0.90, 1.11)	1.01 (0.95, 1.08)	1.02 (0.84, 1.19)	0.98 (0.92, 1.04)	1.00 (0.91, 1.09)	0.97 (0.88, 1.06)	0.97 (0.87, 1.06)
k	**1.16 (1.15, 1.18)**	**0.95 (0.94, 0.96)**	**0.90 (0.89, 0.92)**	**0.93 (0.91, 0.94)**	**0.98 (0.97, 0.99)**	1.00 (0.99, 1.02)	**0.96 (0.95, 0.98)**	1.08 (1.06, 1.09)
R^2	0.64	0.72	0.53	0.44	0.60	0.65	0.46	0.50
SRMSE	2.2	3.1	2.6	4.7	2.4	4.3	3.4	3.0
RNWP	0.8	0.8	1.0	1.2	0.9	1.0	1.1	1.1
n respondents	4977	1001	3307	426	4097	492	1398	1603

Note: 95% confidence intervals for parameter estimates are indicated in parentheses. Bold cells indicate 95% confidence intervals that do not overlap zero for β and α or one for γ and *k*. An odds ratio is obtained for a binary attractiveness variable α using the formula 2^{α}. The attractiveness contribution of a continuous attractiveness variable α with value *x* is obtained by the expression x^{α}.

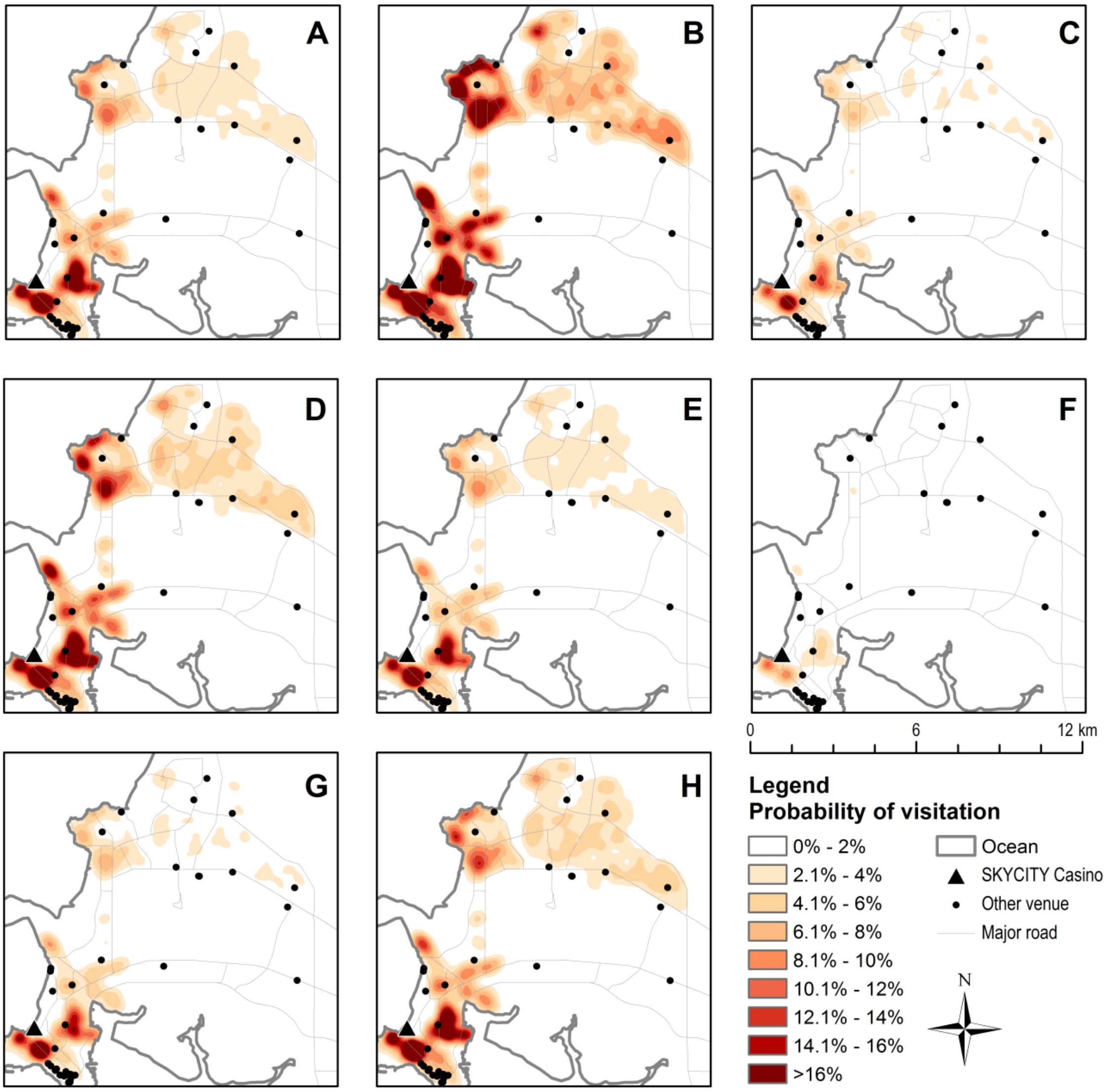

Figure 2.2: Estimated catchments of the SKYCITY Casino, Darwin, among different subgroups. A = all visitors, B = EGM gamblers, C = Non-gamblers, D = moderate to high risk visitors (PGSI ≥ 3), E = non-problem gamblers (PGSI = 0), F = Walkers, G = frequent visitors (≥ 4 visits per month), H = Infrequent visitors (1 visit per month).

Casino's catchment covers the entire residential area of the town. In contrast, the Beachfront Hotel, Darwin, (Figure 2.3) had a moderate sized catchment for all visitors, covering only a few neighbouring suburbs. Similarly, the Casuarina Club's catchment in Darwin (Figure 2.4) was large yet localised to its surrounding region.

Different subgroups are attracted to different venues. Comparing the catchments for EGM gamblers with that for all visitors, SKYCITY Casino catchment (Figure 2.2) is largely unchanged in extent but substantially more intense, indicating a greater probability of EGM gamblers visiting this venue. In contrast, non-gamblers travel further to the Beachfront Hotel (Figure 2.3) than EGM gamblers. Like the casino, the

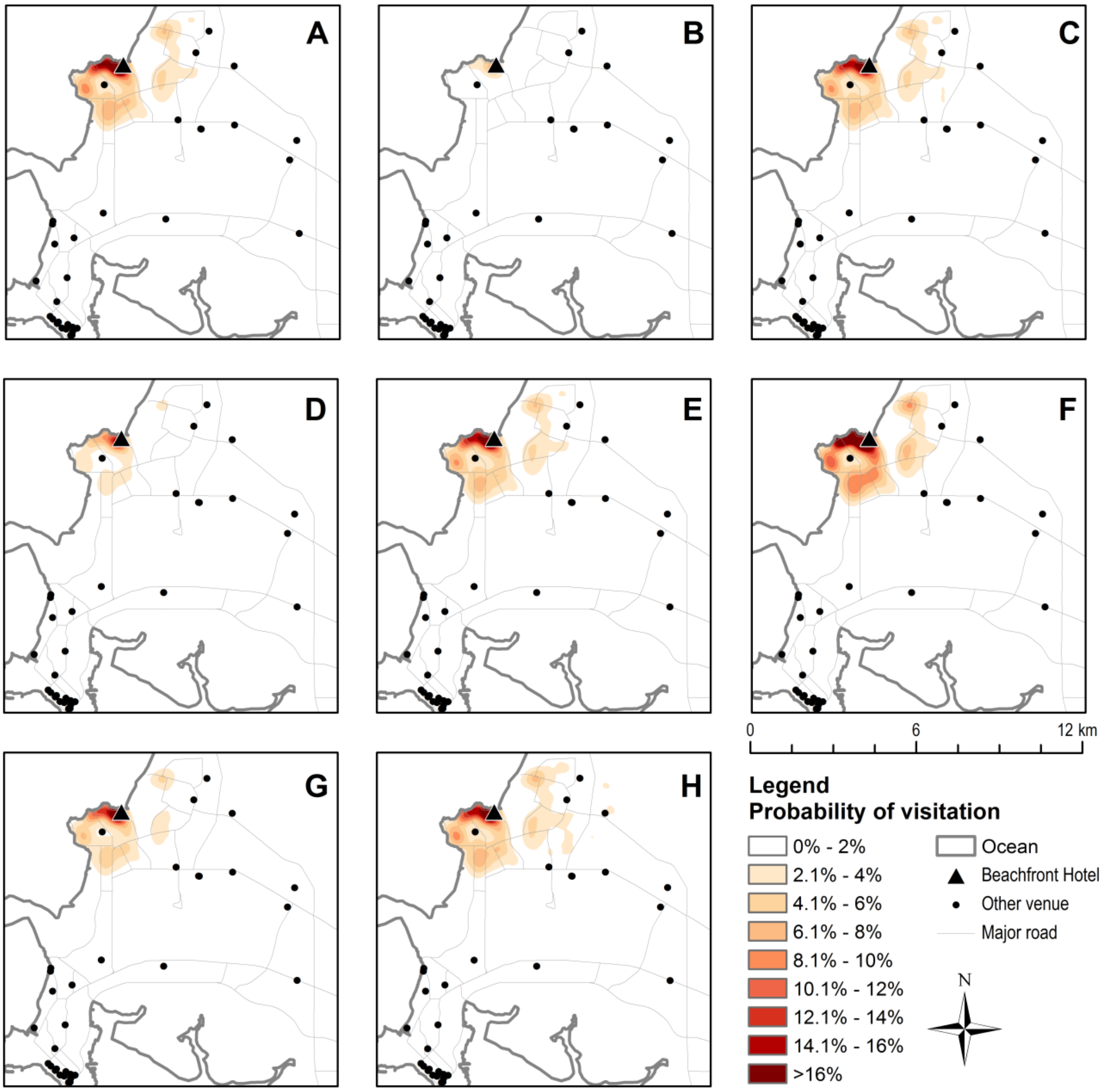

Figure 2.3: Estimated catchments of the Beachfront Hotel, Darwin, among different subgroups. A = all visitors, B = EGM gamblers, C = Non-gamblers, D = moderate to high risk visitors (PGSI ≥ 3), E = non-problem gamblers (PGSI = 0), F = Walkers, G = frequent visitors (≥ 4 visits per month), H = Infrequent visitors (1 visit per month).

Casuarina Club's catchment (Figure 2.4) for EGM gamblers is largely unchanged in size when compared to all visitors although it is slightly more intense.

2.6 Discussion

The calibration process substantially improved the explanatory power of the Huff model, increasing the R^2 from 0.23 to 0.64. The estimated parameters may be usefully employed to predict gambling venue catchments in the NT and other jurisdictions with a similar configuration of EGM supply. This has potential benefits for social impact

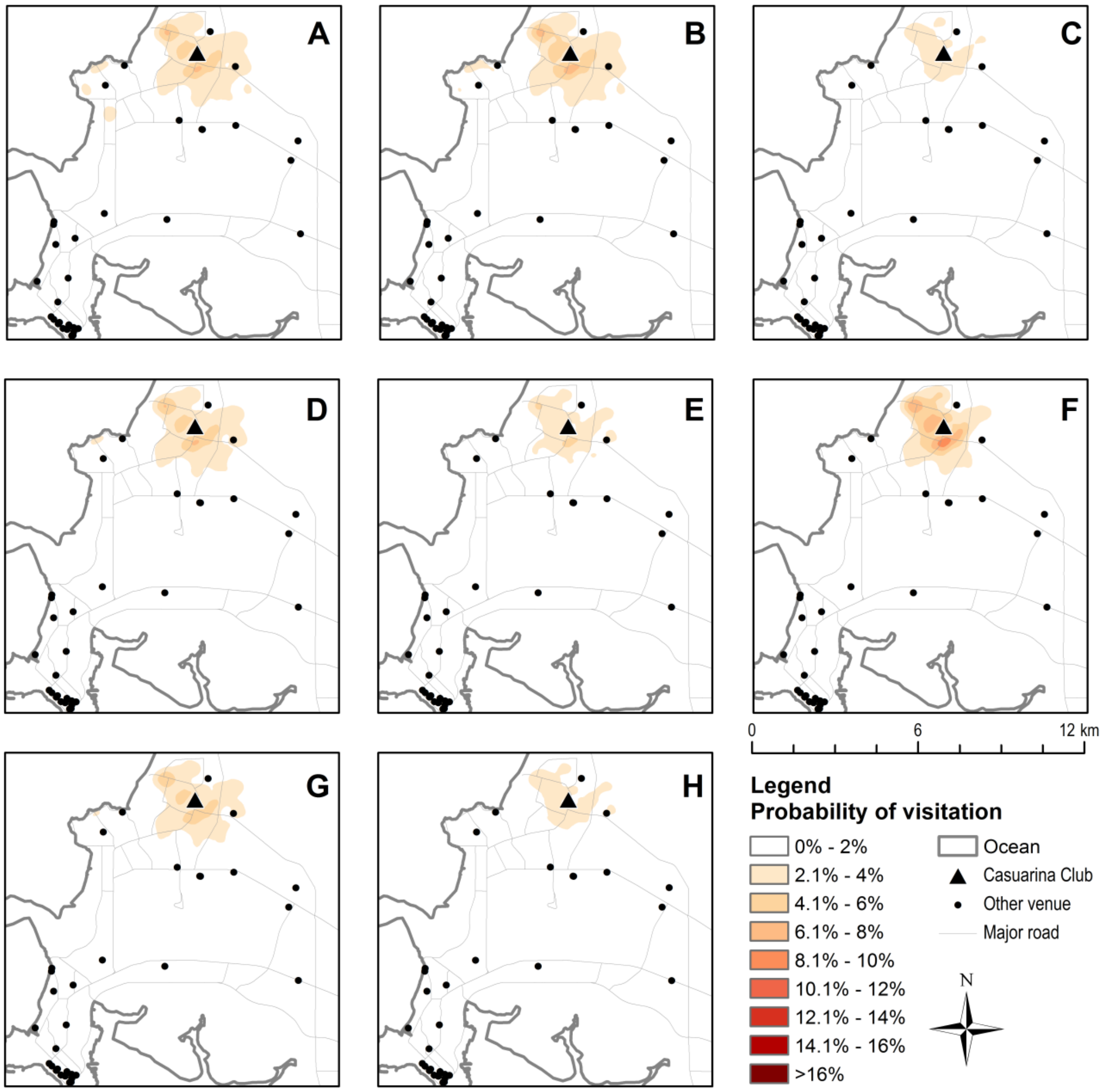

Figure 2.4: Estimated catchments of the Casuarina All Sports Club, Darwin, among different subgroups. A = all visitors, B = EGM gamblers, C = Non-gamblers, D = moderate to high risk visitors (PGSI ≥ 3), E = non-problem gamblers (PGSI = 0), F = Walkers, G = frequent visitors (≥ 4 visits per month), H = Infrequent visitors (1 visit per month).

assessment because catchments can be more reliably estimated in places where surveys have not been conducted. Scenario-based models may be built in other jurisdictions such as Western Australia where EGMs have not yet been introduced outside of the local casino to model the spatial extent of potential social impacts of liberalising EGMs at the local level.

More specifically, we have confirmed the importance of various factors influencing venue choice that have previously only been assumed to be important. In particular, EGM numbers were important in explaining venue visitation behaviour for all subgroups. Indeed, for EGM gamblers, it was the only significant predictor of venue attractiveness. The fact that distance decay was particularly important for EGM and problem gamblers indicates that accessibility is more important to these visitors than non-gamblers.

The question of why venue size is especially important for EGM gamblers is worthy of further research. We can speculate several reasons for this. First, it may simply be that when someone sets out to gambling, larger gambling venues are more prominent in that gambler's mental choice set. In other words, when wanting to gamble, more gambling-oriented venues appear more attractive due to their focused nature on that activity. A second possibility may be the additional gambling-related services provided by larger gambling venues. By virtue of their size, venues with more EGMs generally receive more gambling revenue. These revenues can be directed to attract gamblers with goods and services such as free beverages, courtesy buses, loyalty card programs, larger jackpots and better customer services, all of which may be attractive to those wishing to gamble. Finally, gambling in a larger venue can have a 'contagion effect', meaning that gambling in a large and busy venue may encourage longer play (Rockloff et al., 2011). It is plausible that this more intense gambling behaviour may encourage revisits to the venue to repeat this same experience.

While previous research has shown that catchment sizes vary between venues (Doran et al., 2007), we have extended this result to show that catchment sizes vary between groups of visitors, even for the same venue. Furthermore, the direction of this effect is venue specific. Put another way, different population subgroups are attracted to different venues. There is no single set of venue-level variables that optimally represents the 'attractiveness' construct for all patrons. This implies that future studies should consider the subpopulation of interest carefully prior to selecting the composition of venue-attractiveness indices.

These findings have three important implications for research and regulation. First, the catchment considered should be appropriate to the specific preventative or harm-minimisation strategy under investigation. Interventions to reduce the public health impacts of gambling venues may need to consider catchments for non-gamblers, gamblers and problem gamblers separately. For non-gamblers, strategies such a reduction in exposure or accessibility may be considered, while for gamblers and problem gamblers preventative and harm-minimisation strategies, respectively are needed. In terms of targeting interventions, venue visitation choices should be considered. For example, interventions designed to intercept people experiencing gambling-related harm in order to offer treatment and other assistance should be best targeted toward venues with large numbers of EGMs rather than small venues with a handful of EGMs. Third, given that number of EGM venues is a factor that affects venue choice for EGM gamblers and those who score 3 or more on the PGSI, there is a prima facie case to investigate capping venue size. Further research is required to determine if larger gambling venues are more harmful than smaller ones.

Second, the spatial extent of the social impacts of gambling venues should not be assessed without explicit consideration of local factors. The spatial configuration of EGM supply, the characteristics of venues, and the spatial distribution of their patrons are all factors in determining venue visitation behaviour. Spatial behavioural models such as the Huff model may need to be employed in such future assessments.

Our findings are subject to a number of caveats. First, whilst the use of a mail survey based on the G-NAF holds a number of key advantages with respect to accurately geocoding responses, this technique misses hard-to-reach and mobile sub-populations such as visitors from these towns' hinterlands, transient workers, and other groups unlikely to respond to residential surveys. Second, the relatively low response rate, although typical for surveys of this kind (e.g. Nakaya et al., 2007), means that vulnerable groups in the community may be under-represented in our results. A better approach to data collection might involve venue exit surveys, performed in collaboration between industry, regulators and researchers. Third, the findings may not be generalizable beyond settings with dispersed gambling machines (e.g. Australia, New Zealand, most of Canada and several states of the USA).

Future research might usefully focus on replicating these methods and findings in different geographic contexts. Such studies might also investigate the community-level effects of gambling harm and employ the evidence presented here regarding the attractiveness of gambling venues to better assess the role of accessibility in mediating harm. Future problem gambling prevalence surveys could usefully employ modern address matching technologies and collect household-level spatial data to test these relationships in other contexts.

2.7 Conclusions

Our finding that catchments vary not only between venues, but also between different subgroups of visitors, has important public health implications. Some venues attract more vulnerable visitors than others. Our results suggest that harm minimisation measures could be better targeted if they consider specific combinations of gambling venues and visitor subgroups. Spatial modelling can provide decision support for regulators tasked with approving new license applications or for evaluating the impact of health interventions such as mandated reductions in EGM numbers.

Model calibration resulted in a substantial improvement in model fit relative to previously published studies. A calibrated Huff model is suitable for application in urban contexts as a regulatory tool for social impact assessment and harm minimisation. Our parameter estimates might usefully be applied to improve the identification of vulnerability hotspots in other locales. Given the trend toward gambling liberalisation throughout much of the developed world, the ability better understand the spatial relations between gambling venues and the communities that support them is essential.

Chapter 3: The relationship between player losses and gambling-related harm: evidence from nationally representative cross-sectional surveys in four countries

3.1 Foreword

This chapter presents the first in a series of three studies that sought a better understanding of the relationship between EGM gambling losses – also referred to as EGM gambling expenditure – and risk of gambling problematically. There were two reasons for investigating this relationship in the context of this thesis. First, it was clear that the most reliable and most spatially-detailed data available on gambling behaviour is EGM loss data, which are centrally monitored by regulators for taxation and probity purposes. Because these data are available for individual EGM venues, they provide the most spatially detailed indicator of the geographic distribution of EGM gambling. Second, however, it was unclear from the literature if and how EGM gambling losses might operate as a proxy measure for problem gambling. In particular, because the conventional wisdom in gambling studies – sometimes described as the 'responsible gambling' model of gambling behaviour – asserts that 'safe levels' of gambling are possible, it was arguable that among certain populations, EGM gambling losses were not associated with harmful outcomes.

Consequently, it was decided that if EGM gambling losses were to be used to predict problem gambling prevalence in small areas, then a specific investigation of the relationship between losses and harm was warranted. This chapter presents the first in a series of three studies of this relationship. Because this research has important implications well beyond the production of spatial models of problem gambling, this chapter is written for a broader audience with an interest in the epidemiology of gambling. Nevertheless, the findings of this chapter (and the two that follow it) underwrite the validity of methods used to produce small area estimates in Chapter 6.

This chapter was published as:

Markham, F., Young, M. & Doran, B., 2016. The relationship between player losses and gambling-related harm: evidence from nationally representative cross-sectional surveys in four countries. *Addiction*, 111(2), p.320–330.

3.2 Abstract

Background and Aims

Flaws in previous studies mean that findings of J-shaped risk curves for gambling should be disregarded. The current study aims to estimate the shape of risk curves for gambling losses and risk of gambling-related harm a) for total gambling losses and b) losses disaggregated by gambling activity.

Design

Four cross-sectional surveys.

Setting

Nationally-representative surveys of adults in Australia (1999), Canada (2000), Finland (2011) and Norway (2002).

Participants

10,632 Australian adults, 3,120 Canadian adults, 4,484 persons aged 15-74 in Finland and 5,235 persons aged 15-74 in the Norway.

Measurements

Problem gambling risk was measured using the modified South Oaks Gambling Screen, the NORC DSM Screen for Gambling Problems and the Problem Gambling Severity Index.

Findings

Risk curves for total gambling losses were estimated to be r-shaped in Australia (β losses = 4.7 [95% CI 3.8, 6.5], β losses^2 = -7.6 [95% CI -17.5, -4.5]), Canada (β losses = 2.0 [95% CI 1.3, 3.9], β losses^2 = -3.9 [95% CI -15.4, -2.2]) and Finland (β losses = 3.6 [95% CI 2.5, 7.6], β losses^2 = -4.4 [95% CI -34.9, -2.4]) and linear in Norway (β losses = 1.6 [95% CI 0.6, 3.1], β losses^2 = -2.6 [95% CI -12.6, 1.4]). Risk curves for different gambling activities showed either linear, r-shaped, or non-significant relationships.

Conclusions

Player loss – risk curves for total gambling losses and for different gambling activities are likely to be linear or r-shaped. For total losses and electronic gaming machines, there is no evidence of a threshold below which increasing losses does not increase the risk of harm.

3.3 Introduction

The social and health impacts of gambling primarily result from gamblers losing money (Productivity Commission, 2010). Although problem gambling is frequently conceptualised as a behavioural addiction (American Psychiatric Association, 2013), because the loss of money itself is the proximate cause of many of the harms associated with gambling, it is therefore worthy of investigation in its own right. Yet the relationships between money lost gambling and gambling-related harms have rarely been the specific subject of sustained investigation. The only gambling-related harm to have received substantial scrutiny in relation to money lost on gambling is problem gambling risk, as measured by problem gambling screens and their constituent items in numerous problem gambling prevalence studies (Gainsbury et al., 2014; Wardle et al., 2011; e.g. Welte et al., 2002). Money lost should be of particular interest to policymakers and scholars because loss statistics that are routinely collected by governments at the jurisdictional and venue levels could potentially play an important role in regulation.

One line of research into monetary losses from gambling that has received intermittent scholarly attention has been the establishment of safe consumption guidelines. Taking their cue from alcohol consumption guidelines, a handful of studies have sought to identify “safe” levels of gambling consumption (Currie et al., 2008, 2006; Rockloff, 2012; Stinchfield and Winters, 2001; Weinstock et al., 2007). Using a variety of methods, these studies have sought to define a threshold point for gambling consumption which maximises the differentiation between problem and non-problem gamblers. In a much cited Canadian national study (2006) and its replication in three Canadian provinces (2008), Currie and colleagues reported the existence of J-shaped risk curves, analogous to those long reported for the effect of alcohol consumption on coronary heart disease (e.g. Ronksley et al., 2011). On this basis, receiver operating curve (ROC) analyses found low-risk gambling thresholds at \$500-1000/year in Canada (Currie et al., 2006), and \$1020/year, \$400/year and \$132-\$600/year in Alberta, Ontario and British Columbia, respectively (Currie et al., 2008). Weinstock et al. (2007) performed a similar analysis on a sample of 178 post-treatment gamblers in the United States, finding that a threshold of ≤ 1.9% of monthly income spent gambling was the best cut-point for predicting problem-gambling symptoms. Stinchfield and Winters

(2001) and Rockloff (2012) performed related analyses using gambling frequency as the predictor variable, both finding that time spent gambling was useful for discriminating problem from non-problem gamblers. These authors did not differentiate between consumption of different gambling activities.

Underpinning these studies is an understanding that the relationship between gambling-related harm and gambling consumption is J-shaped. Indeed, the existence of J-shaped consumption – risk curve is assumed in much literature in the field of gambling studies and is crucial to the 'responsible gambling' approach to regulation. For example, in their influential *Reno Model*, Blaszczynski et al. (2004) wrote that the first of six "fundamental assumptions" contained within the responsible gambling framework is that "safe levels of gambling participation are possible". Co-author Howard Shaffer (2005) was more explicit in a later publication, where he claimed that "[g]ambling, like drinking alcohol, displays a 'dose–response' association that reflects hormesis as an underlying process." If the consumption – risk relationship is not J-shaped, then there can be little scientific underpinning to safe consumption limits: a linear relationship, for example, would imply that all consumption increases risk of harm. As Midanik et al. (1996) discuss in the context of alcohol, without a clear threshold below which an individual is at zero risk of harm, guidelines need to consider what amount of absolute risk can be tolerated.

Yet the shape of the relationship between money lost gambling and gambling-related harm has been the subject of remarkably little research. Of the studies cited above, only those led by Currie (2008, 2006) sought to empirically describe the shape of the consumption – risk curve. Currie and colleagues found that the "nature of the relationship between risk level and gambling behaviour is best described as J-shaped" (2006). Unfortunately, this result must be ascribed to a flawed interpretation of an artefact resulting from the survey instrument. The Canadian Community Health Survey 1.2 (CCHS) analysed in that paper collected player loss data using ordinal brackets of increasing width, rather than in exact dollars. These brackets were treated as though they were of equal width in the published plots, reproduced in Figure 3.1, panel A. However, a recoding of these brackets using their midpoints and dropping the final open-ended bracket is strongly suggestive of a linear relationship between gambling losses and harm (see Figure 3.1, panel B). A similar pattern can be detected in the replication study, although the data are noisier (see Figure 3.1, panels C and D). The absence of any apparent J-shape in the corrected curves makes the identification of safe gambling thresholds highly problematic.

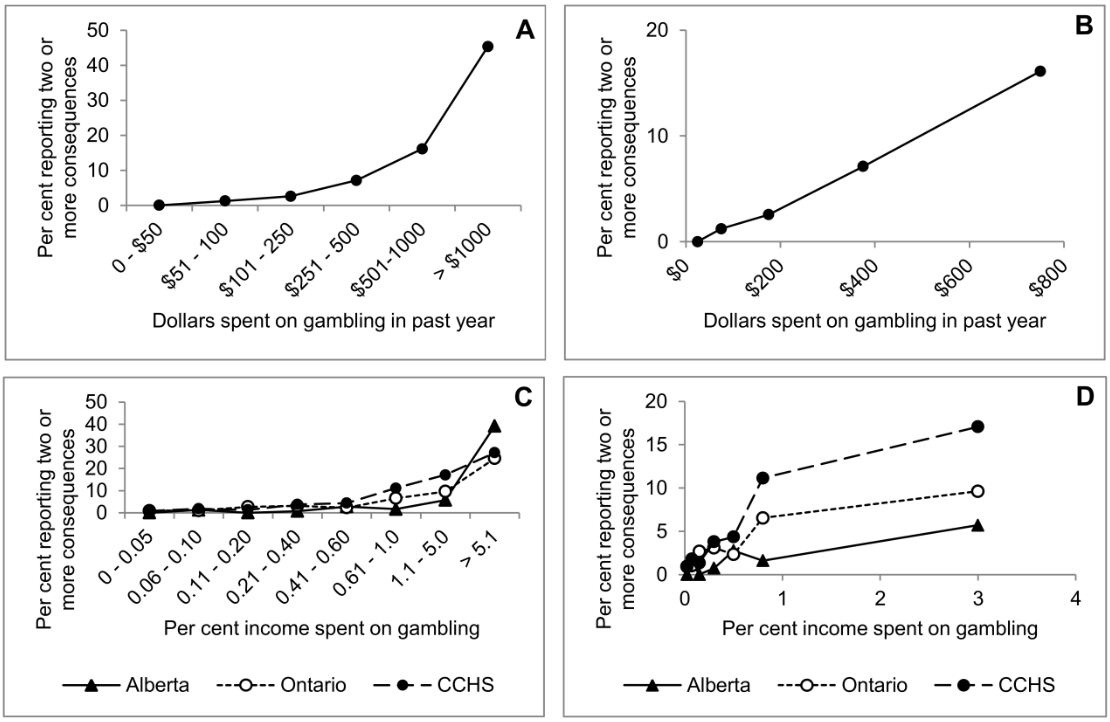

Figure 3.1: When bracketed player loss data are used, the shape of the risk curve depends on how brackets are treated. Panels A and B show the dollars spent on all forms of gambling in the past year by the percentage of respondents reporting two or more negative consequences from gambling. Panels C and D show the percentage of household income spent on all forms of gambling by the percentage of respondents reporting two or more negative consequences from gambling. Panels A and C use the original, non-equal width brackets while panels B and D use midpoint coding and drop the final, unbounded bracket. Data were digitised from Figure 1 in Currie et al. (2006) and Figure 2 in Currie et al. (2008).

A further limitation of published risk curves is that they are presented for total gambling losses only. Many studies have shown that there is great variation in the associations between harm and participation in different gambling activities (e.g. Productivity Commission, 1999; John W. Welte et al., 2004), with recent prospective studies showing that some gambling activities more strongly predict the onset of future harms than others (Billi et al., 2014; el-Guebaly et al., 2015; e.g. Williams et al., 2015). It is likely, therefore, that consumption – risk curves vary between gambling activities.

The purpose of the current study is to identify the shape of the association between monetary gambling losses and problem gambling risk for different gambling activities. To do so, we perform a secondary analysis of four nationally-representative cross-sectional surveys from Australia, Canada, Finland and Norway. Using bootstrapped local polynomial regression, multiple linear regression and mixed effects linear models we:

1. Estimate the shape of gambling loss – problem gambling risk curves for total gambling losses
2. Estimate the shape of gambling loss – problem gambling risk curves disaggregated by gambling activity

3.4 Methods

3.4.1 Data

Player loss data are typically subject to several shortcomings. In particular, non-gambling specific household surveys dramatically underreport gambling losses (Productivity Commission, 2010), although gambling specific surveys also encounter underreporting (e.g. Gerstein et al., 1999). While improvements to sample design and question format have mitigated these problems (Wood and Williams, 2007), it is plausible that survey instruments will impact results. Therefore, we have taken a replication approach to improve confidence in our findings. By using multiple datasets collected with different survey instruments, we hope to determine if our findings are robust across differently collected samples.

Secondary data sets were sought for Australia, Canada, Ireland, Denmark, Finland, New Zealand, Norway, the United States of America, the United Kingdom, Singapore, and Sweden. These countries were selected because, at the time of writing, they were the countries with the highest per capita gambling losses. Using a list of prevalence studies (Williams et al., 2012), we searched for data sets which: were nationally representative; were from a country reporting high levels of gambling losses; included a validated screening test for problem gambling; included questions about gambling expenditure in which losses were recorded as a continuous variable rather than a bracketed ordinal variable; and in which questions about gambling losses were disaggregated by gambling activity. Four studies were identified as suitable using this protocol and were available for reanalysis.

Where appropriate, questions about losses were combined (e.g. questions about gambling losses on lottery tickets purchased online or in stores were combined). To minimise differences in units between jurisdictions and time periods, player loss variables were converted to 2013 currency units using country-specific consumer price indices. These were then converted to US dollars per month using exchange rates adjusted for purchasing power parity for private consumption (The World Bank, 2015). In all studies, socio-demographic questions elicited respondents' sex, age, employment status, education level, income and marital status.

The Australian *National Gambling Survey* was a nationally-representative telephone interview survey of Australian adults, conducted in March and April 1999 (Lattimore and Phillips, 2000). The measure of gambling-related harm in this study was a problem gambling screen, the modified South Oaks Gambling Screen (SOGS-M). The SOGS-M reframes questions from the original SOGS (Lesieur and Blume, 1987) to only enquire about the last 12 months. SOGS-M was only administered to those who gambled at least 52 times per year or whose annual gambling losses reached $4,000 AU1999. While 10,632 people responded to the survey (response rate = 47%), only the 1,240 who completed SOGS-M were included in this study. A complex series of questions were used to elicit information from which losses were calculated. For example, respondents who gambled at racecourses were asked "*Thinking of when you go to a racecourse, how much money do you usually take with you to bet on the races, including any additional money withdrawn or borrowed during your time at the races? And how much money do you usually have left when you leave the races?*" Responses were combined with questions on gambling frequency to estimate annual losses. In the present study, the absolute value of losses was used to simplify estimates, following Welte et al. (2002). For a detailed account of the survey see Productivity Commission (1999).

The *Canadian National Validation Survey* was a telephone interview survey of Canadian adults undertaken between February and April 2000, as part of the development of the Canadian Problem Gambling Index (Ferris et al., 2000). The survey included the Problem Gambling Severity Index (PGSI), a validated problem gambling screen (Ferris and Wynne, 2001b). Losses on individual gambling activities were estimated by asking "*How much money, not including winnings, do you spend on raffle or fundraising tickets in a typical month?*" All respondents were administered the PGSI. 3,120 completed responses were recorded.[3] For a detailed account of the survey see Ferris and Wynne (Ferris and Wynne, 2001a, 2001b).

The Finnish *Gambling Survey 2011* was a representative telephone interview survey of 4,484 people on the Finnish population register aged 15-74, undertaken between October 2011 and January 2012 (response rate = 28%) (National Institute for Health and Welfare, 2013). The PGSI was administered to those who had participated in gambling in the past twelve months (n = 3,451) and was used to measure gambling-related harm in the present study. For each activity for which respondents reported past 12 month gambling participation, respondents were asked questions like "*How much MONEY did you spend on the following in the past 30 days (in euros)? Please include all the money you used regardless of whether you lost or won.*" For a detailed account of the survey see Castrén et al. (2013).

The *Gambling in Norway* 2002 survey was a representative multi-modal survey of 5,235 Norwegians aged 15-74, undertaken in 2012 (response rate = 55%). The last 12 months version of the NORC DSM Screen for Gambling Problems (NODS) was used to measure gambling-related harm in the present study, with the screen administered to all respondents who reported ever gambling. Lifetime non-gamblers assigned a score of zero. For more details regarding NODS, see Gerstein et al. (1999). For each activity for which respondents reported participation, respondents were asked questions like "*Approximately how much money have you gambled for on gambling machines in the last 30 days?*" For a detailed account of the survey see Lund (2006).

The outcome variable used in all analyses was a validated problem gambling scale (SOGS, PGSI or NODS). These scales were treated as a continuous measure of the harm continuum, with increasing scores on the scale indicating elevated levels of harm.

[3] No information regarding response rate is currently available from published sources, the data documentation, the first author of the study or the data collection agency.

3.4.2 Statistical analysis

Player loss – problem-gambling risk curves were visualised using loess, a locally weighted, non-parametric polynomial regression (Cleveland et al., 1992). In order to reduce the influence of outliers and endpoints and to better communicate the variation in the data, an ordinary, non-parametric bootstrap was used to draw 1,000 loess fits for each risk curve. The optimal span parameter for loess fits was selected for each bootstrap draw by minimising AICc (Hurvich et al., 1998). A separate curve was drawn for total player losses and losses by gambling activity for each of the four surveys. Y-axes were adjusted to align the problem gambling scales using the standardised thresholds suggested by Williams and Volberg (2013) for problem gambling: SOGS-M = 4; PGSI = 5; NODS = 3. To emphasise the range of player losses that includes the vast majority of respondents, risk curves were not plotted beyond $2,000 US dollars per month, although loess fits included the full data range. For plots of the region between $0 and $250 US dollars, see Appendix C (Figures C.1 and C.2).

Regression analysis was used to identify the significance of curves identified in the bivariate analysis after adjusting for differential risk among population subgroups. Multiple linear regression was used with problem gambling screen scores as the outcome variable. The variable of interest, player loss on a gambling activity and the square of that value were both included as predictor variables. Covariates used to account for differential risk among demographic groups were: age (including a quadratic term); sex; education level; marital status; employment status; income (including a quadratic term). Due to the presence of influential observations with large gambling losses, estimates were calculated using the ordinary, non-parametric bootstrap with 5,000 replications and 95% confidence intervals approximated using the percentile method. While p-values were not generated from bootstrap estimates, beta coefficients are considered statistically significant at the 0.05 level if the 95% confidence interval does not contain zero.

A bootstrapped mixed effects linear model with random intercepts and random slopes for the loss variables was also estimated for total gambling losses. Differences in activities across countries precluded their inclusion in mixed effects models. This model pooled data from all four surveys, with the survey as the grouping variable, in order to increase the parsimony of the model and improve the precision of estimates.

Population weights were not used in the regression analyses, although their use did not influence results substantively (see Table C.1 in Appendix C). Curve shape can be inferred by interpreting the estimated coefficients in the linear regression analysis. A positive coefficient for the quadratic loss term implies a J-shaped curve while a negative coefficient implies an r-shaped curve. Missing data were removed listwise, so sensitivity analyses were conducted to estimate the potential impact of missing data on results.

All analyses were conducted using R with the *boot* and *lme4* packages. Human ethics approval was granted by the Human Research Ethics Committee of The Australian National University (protocol no. 2014/313).

3.5 Results

Responses from 8,884 individuals who completed the problem gambling screen across the four surveys are summarised by loss tercile in Table 3.1.

Visual inspection of loess curves for total gambling losses in Figure 3.2 suggests a slightly r-shaped curve in all four surveys. As gambling losses mount, so too does the average risk as measured by the various problem gambling screens. There is no low-risk region of the curve where increasing gambling losses does not increase harm. A flattening of the risk gradient is evident in the Canadian data at a mean PGSI level of around 3.

The risk curves for individual gambling activities presented in Figure 3.3 are diverse. For EGMs, the risk curves all appear r-shaped, with particularly steep gradients in Australia and Norway and a truncated arc in Finland where no respondent reported spending over $320 US dollars per month. Lottery risk curves appear much flatter than those for EGMs, with a generally linear appearance. Risk curves for racing varied in shape, with a steep, linear curve in Australia and flatter gradients in other countries. Risk curves for sports betting and table games were noisy and variable between surveys.

Additional analysis using standard, non-bootstrap methods found similar results (Appendix C, Figures C.3 and C.4).

The analyses reported in Table 3.2 largely confirmed these qualitative relationships. Significant negative quadratic loss coefficients were found for total gambling losses in Australia, Canada and Finland, indicating r-shaped curves, while a linear relationship was found for total gambling expenditure in Norway. Statistically significant r-shaped curves were observed for: EGM gambling in Australia, Canada and Norway; lotteries in Canada and Finland; and sports betting in Norway. Statistically significant linear relationships were observed for EGMs and sports betting in Finland; and racing in Australia. The results were inconclusive as to whether or not an association is present for the eleven remaining risk curves. The only gambling products which may entail a J-shaped dose-response relationship were lotteries in Australia and racing in Finland, but estimates of these curves did not reach significance. Gambling losses predicted up to 25% of the variation in problem gambling risk, depending on country and gambling activity, with gambling on EGMs in Australia and Norway showed the strongest evidence for r-shaped curves. However, losses on lotteries and table games in particular explained very little variation in problem gambling scores. The sensitivity analysis for missing data found few substantive changes (Appendix C, Tables C.4 – C.26). The sensitivity analyses suggested that total gambling losses in Australia and lottery losses in Canada may be linear rather than r-shaped, while racing in Canada may have an r-shaped dose-response curve.

3.6 Discussion and conclusions

3.6.1 Key results

The risk curves for total gambling losses showed no evidence of J-shaped relationships between loss and risk. Where previous studies (Currie et al., 2008, 2006) found J-shaped curves, in this study r-shaped and linear curves more accurately describe the loss – risk relationship. Linear regression analysis confirmed these findings, with significant r-shaped curves found for total gambling losses in Australia, Canada and Finland and a linear curve found in Norway. The mixed effects linear model estimates were comparable with the unpooled estimates from multiple linear regression. Furthermore, none of the 20 activity-specific risk curves appeared to be J-shaped. Risk curves either appear to be r-shaped or linear. However, some linear risk curves had flat gradients (e.g. table games in Australia), implying that for these activities, risk is not directly related to the magnitude of player losses. Indeed, considerable variation was evident among risk curves. EGM gambling was the activity at which losses most strongly correlated with harms. Little relationship was found between losses and harm for table games, with relationships varying between countries for lotteries, racing and sports betting.

3.6.2 Limitations and generalisability

These findings are subject to four important limitations. First, due to the format of household income data in survey questionnaires, this study used absolute amounts lost by gamblers as the explanatory variable, rather than proportions of household income. Use of proportions of income may improve face validity. Second, these estimates rely on self-reported player losses. These are likely to underestimate true losses for activities like EGMs (Productivity Commission, 2010; Wood and Williams, 2007). Third, the differences between survey instruments used in these surveys means that risk curves are not strictly comparable between countries. In particular, the use of different questions to estimate gambling losses is likely to impact the gradients of harm-loss curves. Differences in curve shape between countries may in part be due to different inclusion criteria employed in the surveys (e.g. whether weekly or lifetime gamblers were administered the problem gambling screen). Fourth, these risk curves are based on cross-sectional studies. As recent longitudinal studies (e.g. Billi et al., 2014; el-Guebaly et al., 2015; Williams et al., 2015) found that EGM gambling is a strong predictor of the future onset of gambling problems, we conjecture that similar relations may be found prospectively.

The generalisability of specific risk curves across gambling contexts is limited as the socio-technical determinants of gambling risk vary between jurisdictions and over time. For example, the replacement of EGMs with more restricted machines in 2009 in Norway means that the risk curves documented here for EGMs may no longer apply. Caution should therefore be exercised when generalising to other jurisdictions or within the same jurisdiction if the accessibility of gambling products or their characteristics has altered. Nevertheless, if the gambling environment remains constant, we see little reason to expect that the shape of these risk functions vary over time. While the rate of problem gambling has plateaued in some jurisdictions consistent with the ‘adaptation hypothesis’ (Abbott, 2006), so too have per capita gambling losses (e.g. Abbott et al., 2016b).

Table 3.1: Descriptive statistics of variables of interest, disaggregated by tercile of total gambling losses per month.

	Loss tercile 1 n = 2940			Loss tercile 2 n = 2979			Loss tercile 3 n = 2965		
Numerical variables	**mean**	**sd**	**n**	**mean**	**sd**	**n**	**mean**	**sd**	**n**
Problem gambling score [a]	0.03	0.13	2940	0.05	0.20	2979	0.24	0.51	2965
Losses (USD 2013)	8	6	2940	31	10	2979	238	430	2965
Age	45	16	2932	47	15	2967	45	16	2944
Income (USD 2013)	19859	32079	2417	34093	37155	2371	30919	40879	2300
Categorical variables	**n**	**%**		**n**	**%**		**n**	**%**	
Employment	2930			2970			2949		
Employed	1746	60%		1984	67%		1977	67%	
Not employed	1184	40%		986	33%		972	33%	
Education	2921			2956			2937		
School	1305	45%		1649	56%		1705	58%	
Post-school	1616	55%		1307	44%		1232	42%	
Survey	2940			2979			2965		
Australia 1999	214	7%		258	9%		756	25%	
Canada 2000	393	13%		716	24%		966	33%	
Finland 2011	1687	57%		972	33%		738	25%	
Norway 2002	646	22%		1033	35%		505	17%	

Notes: [a] Problem gambling score standardised according to the problem gambling thresholds calculated by Williams & Volberg (2009): SOGS-M = 4; PGSI = 5; NODS = 3.

Table 3.2: Multiple linear regression and mixed effects linear model estimates of player loss – problem gambling risk curves by gambling activity

		Australia 1999 [95% confidence interval]	Canada 2000 [95% confidence interval]	Finland 2011 [95% confidence interval]	Norway 2002 [95% confidence interval]
Total	$10^3 \times \beta$ *losses*	**4.7 [3.8, 6.5]**	**2.0 [1.3, 3.9]**	**3.6 [2.5, 7.6]**	**1.6 [0.6, 3.1]**
	$10^7 \times \beta$ *losses*^{2}	**-7.6 [-17.5, -4.5]**	**-3.9 [-15.4, -2.2]**	**-4.4 [-34.9, -2.4]**	-2.6 [-12.6, 1.4]
	losses R^2	0.24 [0.18, 0.32]	0.06 [0.03, 0.12]	0.10 [0.07, 0.17]	0.14 [0.07, 0.27]
	n	896	1259	3004	1875
Total (mixed effects)	$10^3 \times \beta$ *losses*	**4.6 [3.6, 6.4]**	**1.9 [1.3, 3.7]**	**3.8 [2.6, 7.7]**	**1.7 [0.6, 2.9]**
	$10^7 \times \beta$ *losses*^{2}	**-7.3 [-18.1, -4.3]**	**-3.3 [-13.9, -2.1]**	**-4.6 [-34.7, -2.6]**	-2.8 [-10.9, 1.2]
	losses variance explained	0.15 [0.06, 0.29]			
	n	7034			
EGMs	$10^3 \times \beta$ *losses*	**6.4 [5.2, 10.5]**	**3.3 [1.3, 8.3]**	**38.3 [23.2, 51.6]**	**5.5 [2.9, 20.8]**
	$10^7 \times \beta$ *losses*^{2}	**-13.2 [-45.3, -9.2]**	**-8.3 [-38.7, -2.0]**	-627.6 [-1207.5, 118.0]	**-9.9 [-322.2, -2.1]**
	losses R^2	0.26 [0.19, 0.36]	0.04 [0.01, 0.18]	0.20 [0.05, 0.32]	0.23 [0.15, 0.52]
	n	619	462	1156	180
Lotteries	$10^3 \times \beta$ *losses*	1.1 [-5.5, 6.1]	**12.2 [4.3, 20.4]**	**6.2 [4.2, 11.0]**	3.1 [-1.9, 5.6]
	$10^7 \times \beta$ *losses*^{2}	35.8 [-67.1, 193.7]	**-450.2 [-846.5, 12.6]**	**-43.4 [-231.4, -28.6]**	-30.8 [-59.5, 314.3]
	losses R^2	0.02 [0.00, 0.06]	0.02 [0.01, 0.05]	0.04 [0.01, 0.10]	0.03 [0.00, 0.10]
	n	722	1073	2700	1943
Racing	$10^3 \times \beta$ *losses*	**2.7 [0.7, 5.5]**	7.9 [-7.7, 43.8]	-1.1 [-3.9, 5.8]	0.8 [-0.9, 8.2]
	$10^7 \times \beta$ *losses*^{2}	-1.7 [-17.5, 16.2]	-200.9 [-1686.9, 287.0]	6.6 [-53.0, 16.0]	-3.4 [-63.8, 5.6]
	losses R^2	0.05 [0.00, 0.16]	0.00 [0.00, 0.76]	0.10 [0.00, 0.36]	0.11 [0.03, 0.60]

	n	453	68	215	101
Sports betting	$10^3 \times \beta$ *losses*	-0.9 [-28.0, 14.3]	8.0 [-7.1, 17.3]	**9.7 [3.7, 15.5]**	**4.5 [2.1, 10.2]**
	$10^7 \times \beta$ *losses* ^{2}	-51.5 [-417.0, 1618.9]	-58.3 [-156.7, 450.1]	-49.2 [-225.4, 32.3]	**-28.3 [-134.5, -11.5]**
	losses R^2	0.00 [0.00, 0.24]	0.08 [0.05, 0.26]	0.06 [0.00, 0.21]	0.15 [0.08, 0.29]
	n	175	222	435	402
Table games	$10^3 \times \beta$ *losses*	4.1 [-3.4, 10.2]	1.0 [-5.9, 3.8]	30.1 [-11.6, 51.0]	0.0 [-0.9, 2.2] †
	$10^7 \times \beta$ *losses* ^{2}	-28.3 [-114.6, 66.5]	-4.9 [-25.8, 80.3]	-225.9 [-1023.7, 3783.1]	0.3 [-10.8, 3.8] †
	losses R^2	0.00 [0.00, 0.16]	0.00 [0.00, 0.24]	0.12 [0.00, 0.45]	0.21 [0.00, 0.90] †
	n	169	126	172	31

Notes: Player loss β coefficients estimated from multiple linear regression or mixed effects linear models. Square brackets report 95% confidence intervals, estimated via the percentile method from an ordinary, non-parametric bootstrap with 5,000 replications. Estimated coefficients are not reported for socio-demographic predictor variables for reasons of brevity. Non-reported predictor variables include: age; age^2; sex; education level; marital status; employment status; household income; and household income^2. *Losses* R^2 reports the variance explained by the player loss terms in the regression, after adjusting for other covariates. *Losses* R^2 was calculated by subtracting the adjusted R^2 of the full multiple linear regression from that of a multiple linear regression specified identically except with the player loss terms dropped. † indicates that due to the small number of observations, the linear regression was specified without socio-demographic predictor variables.

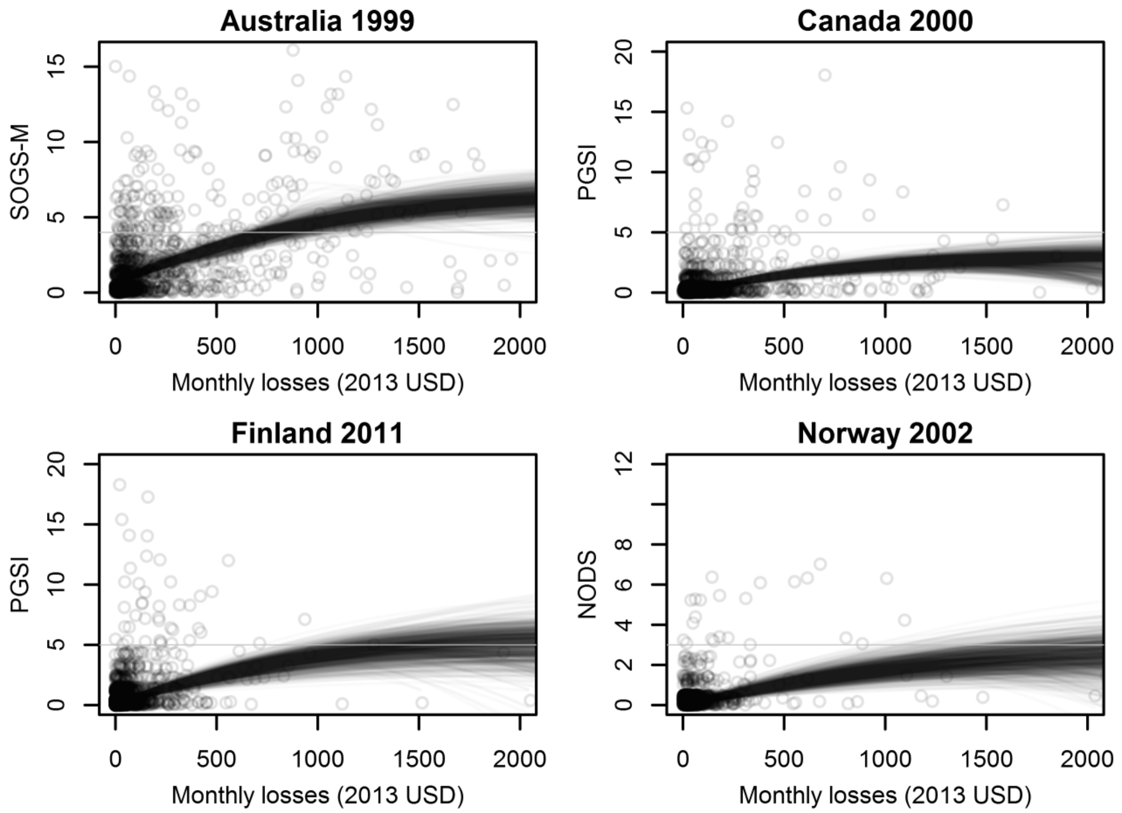

Figure 3.2: Bootstrapped risk curves for total gambling losses versus problem gambling risk. Horizontal lines represent the standardised problem gambling thresholds calculated by Williams & Volberg (2009): SOGS-M = 4; PGSI = 5; NODS = 3. Losses are standardised to 2013 US dollars spent in previous 30 days. Each point represents a single respondent, jittered for display. Each line represents a single non-parametric bootstrapped loess fit, with span selected by AICc. Median spans [and 95% CIs] were: 1.0 [0.4, 5.0], 1.0 [0.8, 5.0], 1.0 [0.6, 5.0] and 1.4 [0.6, 5.0].

It may be suggested that r-shaped risk curves are incompatible with the well-known finding that problem gamblers account for a very large proportion of gambling losses . However, simulation results presented in Appendix C.2 in the online supplementary material shows that an r-shaped curve is consistent with a disproportionate problem-gambler loss share.

3.6.3 Implications and conclusions

There is little evidence supporting the hypothesis of J-shaped risk curves for total gambling losses. Previous studies showing J-shaped curves are methodologically flawed. Risk curves for total gambling losses are likely to be linear or r-shaped. This does not mean that there are no individuals who gamble large amounts of money without experiencing harms. Rather, every increase in consumption increases the risk of harm. In consequence, previous recommendations (Currie et al., 2008, e.g. 2006)

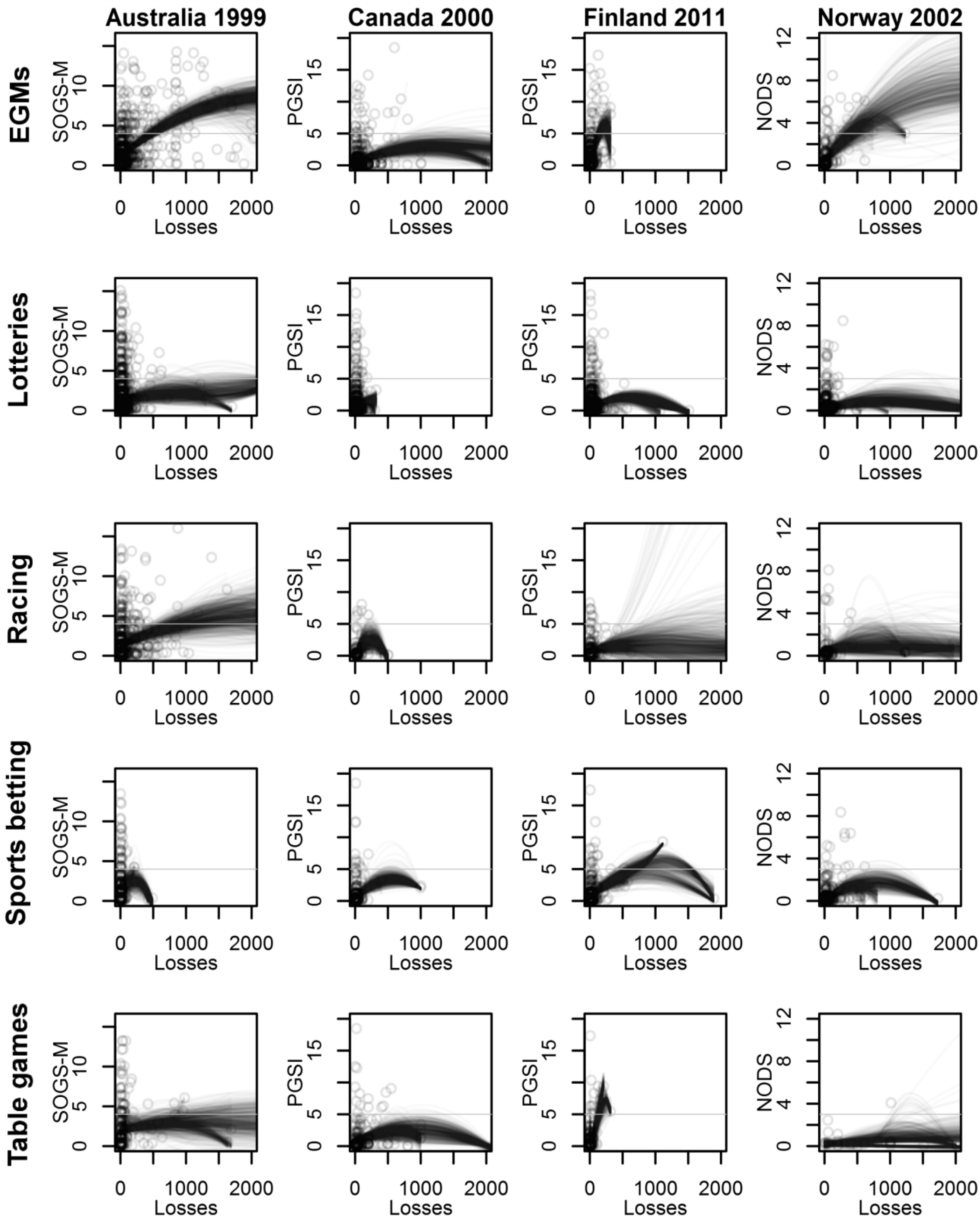

Figure 3.3: Bootstrapped risk curves for gambling losses versus problem gambling risk for five gambling activities. Horizontal lines represent the standardised problem gambling thresholds calculated by Williams & Volberg (2013): SOGS-M = 4; PGSI = 5; NODS = 3. Losses are standardised to 2013 US dollars spent in previous 30 days. Each point represents a single respondent, jittered for display only. Each line represents a single non-parametric bootstrapped loess fit, with span selected by AICc. Median spans from left to right, top to bottom [and 95% CIs] were EGMs: 2.8 [0.9, 5.0], 1.0 [0.7, 5.0], 1.0 [0.5, 5.0], 1.2 [0.8, 5.0]; Lotteries: 2.5 [0.7, 5.0], 5.0 [0.8, 5.0], 1.7 [0.6, 5.0], 1.0 [0.7, 5.0]; Racing: 1.3 [0.7, 5.0], 5.0 [0.8, 5.0], 1.2 [0.6, 5.0], 2.4 [1.1, 5.0]: Sports betting: 1.5 [1.0, 5.0], 5.0 [1.1, 4.9], 1.3 [0.8, 5.0], 1.6 [1.0, 5.0]; Table games: 1.7 [0.6, 5.0], 5.0 [0.9, 5.0], 1.0 [0.6, 5.0], 2.4 [0.9, 5.0].

regarding “safe” levels of gambling should be disregarded and future guidelines must be made on the basis of tolerable levels of risk. Where r-shaped curves are found, risks escalate most quickly per dollar for the initial dollars lost. After a certain sum of money is lost, increased losses appear to have reduced impact on the marginal risk of harm. It is probable that curve shapes depend on the type of harm examined and the instrument by which it is measured.

As previous studies demonstrate (e.g. John W. Welte et al., 2004), different gambling activities appear to be associated with different risk functions. Some gambling activities have only a negligible association with harm, while EGMs exhibit a strong loss – harm relationship, stronger than that for total gambling losses. These relations appear to be moderated by national context. As such, one-size-fits-all consumption guidelines across gambling activities are likely to be inappropriate.

These findings have implications for the ‘responsible gambling’ model of regulation. Contrary to Shaffer’s (2005) assertion that gambling entails a hormesis relationship, many gambling products appear to be more similar to tobacco than to alcohol, in that there is no threshold below which consumption does not increase risk. For EGMs in particular, every increase in consumption increases the risk of harm.

Chapter 4: Gambling expenditure predicts harm: Evidence from a venue-level study

4.1 Foreword

This chapter undertakes a further study of the relationship between EGM losses and problem gambling symptoms, this time at the spatial scale of the EGM venue. This is the scale that is perhaps the most relevant for the social impact assessment process, as venues are the entities that trigger an impact assessment by submitting applications to operate EGMs. Put simply, this chapter asks whether more profitable EGM venues are also more harmful.

This study builds on both the previous chapters. It uses the Huff model calibrated in Chapter 2 to allocate residents to particular EGM venues (or equivalently, to distribute EGM losses to residential areas). It builds on the finding of Chapter 3 (that the dose-response relationship between EGM losses and problem gambling is linear or r-shaped), testing for a possible ecological fallacy. Specifically, although it is clear that more expenditure increases the risk of harm for individuals, it is does not follow that more expenditure in venues must result in more harmful venues. This could occur, for example, if a particular venue attracted a very high-income clientele for whom losing large amounts of money gambling was not an indicator of problem gambling. Consequently, a further study was required to test for the existence of a harm-expenditure relationship at the venue level.

Although this venue-level study follows logically from the individual-level study presented in Chapter 3, it makes no reference to the individual-level study because it was published prior to it.

This chapter was published as:

Markham, F., Young, M. & Doran, B., 2014. Gambling expenditure predicts harm: evidence from a venue-level study. *Addiction*, 109(9), p.1509–1516.

4.2 Abstract

Background and Aims

The Total Consumption Theory of gambling suggests that gambling expenditure is positively associated with gambling-related harm. We test the hypothesis that electronic gaming machine (EGM) expenditure predicts gambling-related harm at the level of the EGM venue.

Design and Setting

Cross-sectional analysis of survey and administrative data relating to the general urban adult population of the Northern Territory of Australia.

Participants

Sample consisted of 7049 respondents to a mail-survey about venue visitation and gambling behaviour across 62 EGM venues.

Measurements

Gambling-related harm was defined as the endorsement of two or more items on the Problem Gambling Severity Index. We obtained venue-level EGM expenditure data from the local licensing authority for all venues in the study area. We compared the prevalence of gambling-related harm among patrons aggregated at the venue level with the estimated mean EGM expenditure for each adult resident in the venue's service area using a Huff model, correlation analysis and multivariate binomial regression.

Findings

Aggregated to the venue level (n = 62), per capita EGM expenditure was significantly correlated with rates of gambling-related harm [r = 0.27, n = 62, p = 0.03]. After adjusting for venue type and number of EGMs, an increase in mean per capita monthly EGM expenditure from AUD10 to AUD150 was associated with a doubling in the prevalence of gambling-related harm from 9% (95% CI 6% - 12%) to 18% (95% CI 13% - 23%).

Conclusions

As suggested by the Total Consumption Theory of gambling, aggregate patron electronic gaming machine expenditure predicts the prevalence of gambling-related harm at the venue level.

4.3 Introduction

Estimates of gambling-related harm, particularly via problem gambling prevalence surveys, are costly and time-consuming to produce. Prevalence surveys, because they are based on self-reported behaviour, also tend to underestimate both gambling expenditure (Productivity Commission, 1999; Wood and Williams, 2007) and rates of problem gambling (Productivity Commission, 1999; Williams et al., 2012). Furthermore, prevalence studies tend to adopt different methods, making comparisons problematic even within the same jurisdiction over time (Young, 2013). They also tend to be of insufficient statistical power to detect small changes over time or to investigate the spatial distribution of harms across small areas (Bunkle and Lepper, 2004).

In contrast, detailed gambling expenditure data at the venue level are routinely collected in all developed countries that levy gambling-specific taxes. For example, the Victorian Government, Australia, publicly release data on all gambling venues within the state, including annual electronic gaming machine (EGM) expenditure, venue location and administrative classification (State Government of Victoria, 2014). These administrative data provide an accurate, complete, and consistent longitudinal measure of commercial gambling behaviour at the venue level. However, in the absence of a demonstrated link between gambling expenditure and the prevalence of gambling-related harm, researchers and regulators have been unable to draw inferences about the distribution of harm using gambling expenditure data. If a definite relationship between expenditure and harm can be established, the extant expenditure data may potentially be used to estimate changes in gambling-related harm over time, and at a fine geographical scale, without the need for expensive and ultimately unreliable prevalence studies.

4.3.1 Literature review

The Total Consumption Theory of gambling, borrowed from the single distribution theory of alcohol studies (Babor et al., 2010; Bruun et al., 1975), implies that the number of people experiencing severe gambling-related harm is correlated with the mean population consumption of gambling (Lund, 2008; Rose and Day, 1990). At the individual venue level, this suggests that the proportion of patrons experiencing severe gambling-related harm is correlated with aggregate gambling expenditure. Similarly, venues with relatively high levels of gambling expenditure per patron will also have relatively high levels of harm. If this proposition is correct, researchers and regulators alike may be justified in using measures of gambling expenditure as a proxy for gambling-related harm within gambling venues.

Most studies examining gambling harm and expenditure have most frequently focused on the individual as the unit of analysis. For example, a nationally-representative study of Canadian adults that specifically examined the relationship between expenditure and harm found gambling expenditure to be a strong predictor of harm (Currie et al., 2006). Unsurprisingly, significant relationships between problem or pathological gambling and gambling expenditure are also consistently found in nationally representative surveys, for example in the United States, Great Britain, Australia, and Sweden (Orford et al., 2013; Productivity Commission, 1999; Rönnberg et al., 1999; Welte et al., 2002).

These correlations at the level of the individual aside, Total Consumption Theory is more concerned with the behaviour of populations. At the regional scale of analysis, a case study of the introduction of the UK national lottery found the mean level of gambling expenditure to be correlated with the number of households spending an excessive proportion of their income on gambling (Grun and McKeigue, 2000). Williams and Wood used secondary data collected in eight Canadian provinces to estimate that problem gamblers (4.2% of the population) accounted for 23.1% of total gambling expenditure (Williams and Wood, 2004). Similarly, Livingstone and Woolley presented data that demonstrated the within-session expenditure of problem gamblers in Victoria was three times that of non-problem gamblers (Livingstone and Woolley, 2007). Hansen and Rossow, in a study of 11,637 adolescents across 73 Norwegian schools found that the school-level prevalence of problem gambling was associated with the mean gambling expenditure among students (Hansen and Rossow, 2008). Room *et al*. found that both the mean level of gambling expenditure and the prevalence of gambling problems increased in the local community after the opening of a casino at Niagara Falls (Room et al., 1999).

With the jurisdiction as the unit of analysis, the Australian Productivity Commission compared rates of problem gambling with EGM expenditure and demonstrated a positive correlation between EGM expenditure and rates of problem gambling in eight Australian states and territories (Productivity Commission, 1999). Similarly, a meta-analysis of 34 problem gambling surveys conducted in Australia and New Zealand since 1991 found a strong, positive relationship between problem gambling prevalence and the per capita density of EGMs, although expenditure was not specifically examined in this analysis (Storer et al., 2009).

However, a number of studies have failed to produce clear evidence of a correlation between gambling expenditure and gambling-related harm. As noted by Abbott (2006), the results of a large, national general population survey in the United States were not consistent with the hypothesised relationship between expenditure and gambling harm at the regional level (Welte et al., 2002). Similarly, in several countries, most notably New Zealand, population problem gambling prevalence as estimated by successive surveys has not risen, while aggregate gambling expenditure over the same period had increased substantially (Abbott, 2006).

No study to date has explicitly examined the relationship between gambling expenditure and the prevalence of gambling-related harm *at the venue level*. There are two reasons why the gambling venue level is a particularly important scale for the analysis of gambling-related harm. First, as the site at which most gambling actually occurs in developed countries, regulated gambling venues provide arguably the most important location at which harm minimisation interventions can be targeted. Levels of harm among patrons varies between venues (Clarke et al., 2010; Young et al., 2012b), suggesting that venue-specific factors may play a substantial role in mediating the riskiness of gambling. Second, an emerging body of literature has documented a relationship between heightened problem gambling risk and residential distance to gambling venues at the level of the individual gambler (Pearce et al., 2008; John W Welte et al., 2004; Young et al., 2012a). Yet the causal mechanism which generates an association between proximity to gambling venues and gambling-related harm remains unclear.

If a link can be established between gambling expenditure and gambling-related harm at the venue level, it may advance our understanding of the spatial patterning of gambling-related harm. This study is the first to test the hypothesis that EGM expenditure is correlated with gambling-related harm at the venue level. Furthermore, it describes the strength of that relationship in order to gauge the potential use of per capita EGM expenditure as a predictor of gambling-related harm.

4.4 Methods

4.4.1 Data

To investigate the relationship between gambling expenditure and the prevalence of gambling-related harm at the EGM venue level, three independent sets of data are required: A) estimates of the prevalence of gambling-related harm among patrons of individual venues, B) venue-specific EGM expenditure data, and C) estimates of the number of adults in the service area of each venue, to use as the denominator for estimating per capita EGM expenditure.

4.4.1.1 Gambling-related harm

We obtained venue-level estimates of gambling-related harm by conducting a postal survey. Using the Australian geocoded national address file (G-NAF: PSMA Australia, 2010) as a sample frame, we mailed a questionnaire to all 46,263 households in the urban centres of the Northern Territory to which Australia Post would deliver unsolicited mail and which were zoned residential. To extend our spatial coverage, we selected 2,300 addresses across the peri-urban fringes of the two largest urban centres (to which Australia Post does not deliver mail) for hand delivery of questionnaires. The questionnaires were mailed out once to each address between April and August 2010 and hand delivered in July and September 2010. Any household member aged eighteen or older was eligible to respond, and return of the survey implied consent. The Human Research Ethics Committee of Charles Darwin University granted approval to conduct the study (protocol no. H09048).

To mitigate survey non-response bias we weighted responses using post-stratification. We used raking to estimate weights for the follow strata: gender, age bracket (18-29, 30-44, 45-64, ≥65), town and delivery method (postal- or hand-delivery). We derived strata populations from the profiles of those who were present in the study area on census night during the 2011 Census of Population and Housing.

The questionnaire elicited information about which gambling venues the respondent had visited in the last month. Respondents selected their most frequently visited venue from a list of all EGM venues in, or proximate to, their town of residence. Participants were asked to report whether they participated in EGM gambling on their last visit to this venue and to complete Problem Gambling Severity Index (PGSI: Ferris and Wynne, 2001a) for the last twelve-months. Following Currie *et al.* (2006), we coded those respondents who endorsed two of the nine questions in the PGSI as ‘Sometimes’, ‘Most of the Time’ or ‘Almost Always’ as experiencing gambling-related harm (note that a subsequent analysis of the same dataset using the more conventional categorisation of those scoring 8 or more on the PGSI as the outcome variable yielded similar results in terms of significance but with a larger estimated coefficient for per adult expenditure). The Currie *et al.* measure of gambling harm was selected in order to better capture ‘gambling-related harm’, which is conceptually broader than the pathological gambling construct upon which the conventional PGSI 8+ threshold is based (Currie et al., 2006).

We estimated the prevalence of gambling-related harm for each venue in the study by allocating individual respondents to the venue they had visited most frequently in the previous month. Respondents who did not visit a venue in the last month or who did not complete the PGSI (n = 2,102) were excluded from the analysis.

4.4.1.2 EGM expenditure

We obtained EGM expenditure data for each venue in the study from the state regulatory authority, the NT Department of Justice. This dataset contained nominal monthly EGM expenditure, the number of EGMs operational at the end of each month, the street address and the licensing category (i.e. hotel, club or casino) for each venue in the study. Rather than directly use monthly figures for expenditure and operational EGMs, we adjusted the expenditure series for inflation into September 2010 Australian dollars (AUD) and calculated the mean for both of these series over the period of the survey (April to September 2010).

4.4.1.3 Estimated service-area adult population

We estimated the service-area population of each gambling venue using the Huff model, a probabilistic method for calculating trading areas and their populations (Huff, 1963). We parameterised the Huff model using coefficients derived from a previous analysis of EGM gamblers’ visitation patterns based on the postal survey (Markham et al., 2014b). We used G-NAF dwellings as origin points, weighted according to the adult (aged 18+)

population distribution at the Statistical Area 1 level as counted in the 2011 census. To capture EGM use by non-residents, we used the place of enumeration census dataset, which counts the number of people who were present in a location on census night, as our weighting datum. The study area was defined as all dwellings within 40 km of venues in the study, on the basis that journeys of 40 km or more are generally categorised as irregular rather than commuter trips in Australia (Barry, 1999). The Huff model used took the following form:

$$servicePop_i = \sum_j 0.95 \cdot o_j^{1.01} \cdot \frac{a_i \cdot d_{ij}^{-1.18}}{\sum_i [a_i \cdot d_{ij}^{-1.18}]}$$

where *servicePop*$_i$ is the census-night population of the service area of venue i, o_j is the estimated population of dwelling j, d_{ij} is the Euclidian distance between dwelling j and venue i, and a_i is an index of the relative attractiveness of venue i, defined as:

$$a_i = numEGMs_i^{1.17} \cdot isCasino_i^{-0.23} \cdot isClub_i^{0.12} \cdot \ln supermarketDist_i^{-0.31} \cdot \ln gpoDist_i^{0.26} \cdot ocean_i^{0.2} \cdot innerCity_i^{-0.18}$$

For details regarding these measures, the derivation of their weightings, and more information regarding the service-area model for gambling, see Markham *et al* (2014b).

Descriptive statistics for EGM venues are reported in Table 4.1.

4.4.2 Statistical analysis

We first calculated the Pearson's product-moment correlation between per capita EGM expenditure and the prevalence of gambling-related harm, weighted by the number of responses per venue. We then calculated the association between per capita EGM expenditure and the prevalence of gambling-related harm using a binomial rate regression, an extension of the logistic regression model which analyses the result of multiple Bernoulli trials for each unit (in this case, EGM venues) as the outcome variable. Binomial rate regression was selected as it weights each venue in the analysis according to the number of post-stratification weighted responses, thereby ameliorating the small number problem where rates of gambling-related harm in venues with few survey responses have a much greater variance than those with many responses. As we suspected non-constant variance in regression residuals, we calculated all reported standard errors and confidence intervals using MacKinnon and White's

Table 4.1: Selected medians for gambling venues in the study. Median absolute deviations are reported in parentheses

	Hotels (*n* = 35)	Clubs (*n* = 25)	Casinos (*n* = 2)
Respondents per venue (unweighted)	28 (25)	62 (65)	533 (406)
Respondents per venue (population weighted)	500 (507)	968 (1085)	7803 (5910)
Number of EGMs	10 (0)	22 (18)	531 (354)
Monthly EGM expenditure in AUD	43,253 (23,526)	62,799 (87,370)	3,581,380 (2,557,500)
Harm rate [a]	8.3% (4.7%)	14.6% (5.6%)	19.6% (3.5%)
Service population	444 (78)	1,884 (1,677)	30,812 (26,824)
Monthly EGM expenditure per adult	96 (31)	40 (34)	127 (28)

Note: As most variables are not normally distributed, medians and median absolute deviations are reported instead of means and standard deviations.
[a] The harm rate is the weighted mean of the harm rates of all venues. The weightings were derived from the post-stratification estimates of the number of people in the sample frame who visit that venue most frequently.

heteroskedasticity-correcting estimator (MacKinnon and White, 1985). We calculated the predictor variable of interest, per capita EGM expenditure, by dividing EGM expenditure by the estimated adult service population for each venue. We included other licensing variables, such as venue type (i.e. hotel, club or casino) and the number of operational EGMs, as covariates as previous studies have shown these to be associated with rates of gambling-related harm (Young et al., 2012b). All statistical analyses were determined prior to commencing analysis except for post-stratification weighting, which was conducted following the suggestion of an anonymous reviewer.

4.5 Results

We received 7,049 completed questionnaires, constituting a response rate of 14.5%. As Table 4.2 demonstrates, respondents were older [Wilcoxon rank sum test: W = 53976961, $p < 0.001$], more likely to be female [$\chi^2 = 370.4$, $df = 1$, $p < 0.001$] and better educated [$\chi^2 = 1429.8$, $df = 2$, $p < 0.001$] than the general population (see Table 4.2).

Monthly EGM expenditure per capita and the prevalence of gambling-related harm were significantly correlated at the venue level [$r = 0.27$, $n = 62$, $p = 0.03$] in a bivariate comparison. After fitting the multivariate binomial regression model that controls for the number of EGMs in the venue and the licensing category of the venue (i.e. hotel, club or casino), there was still strong evidence for this correlation (see Table 4.3), a result strengthened by changes to the venue weighting scheme (see Appendix D, Table D.1).

The prevalence of gambling-related harm at a club with the median 22 EGMs is estimated to increase from 9% (95% c.i. 6% - 12%) to 18% (95% c.i. 13% - 23%) as the monthly EGM expenditure per adult rises from AUD10 to AUD150 (see Figure 4.1). In other words, within this range of expenditure (which includes 89% of the venues in the study and 92% of the respondents who visited a venue), each AUD20 increase in monthly EGM expenditure per adult is associated with an estimated average 1.7% increase in the prevalence of gambling harm. Compared to a null model, around 25% of the deviance in the rates of gambling-related harm among patrons was explained by the multivariate binomial regression model. The mean respondent-weighted absolute value of venue residuals was 4.6% (*SD* = 4.0%).

Table 4.2: Demographic composition of sample

	Sample	Population
Sex		
Female	4,300 (62%)	54,351 (50%)
Male	2,652 (38%)	54,476 (50%)
Age		
18-29 years	656 (10%)	26,656 (24%)
30-44 years	1,914 (28%)	33,852 (31%)
45-64 years	3,304 (48%)	36,767 (34%)
65 years or older	971 (14%)	11,552 (11%)
Education level		
School	2,409 (34%)	34,826 (40%)
Tech	1,298 (19%)	29,438 (33%)
University	3,301 (47%)	23,629 (27%)
Employment status		
Self-employed	582 (8%)	8,171 (9%)
Employee	4,827 (69%)	62,441 (66%)
Not in labour force	1,294 (19%)	20,966 (22%)
Unemployed	273 (4%)	2,413 (3%)

Table 4.3: Predictors of the prevalence of gambling-harm in EGM venues

	Coefficient estimate (95% confidence interval)	*p* value
Intercept	-3.15 (-3.98, -2.32)	< 0.0001
Monthly expenditure per adult, 100s AUD	0.58 (0.10, 1.05)	0.0172
Venue type		
Casino	0.00 (ref. group)	
Club	0.74 (0.28, 1.20)	0.0016
Hotel	0.33 (-0.09, 0.74)	0.1287
Number of EGMs, 10s	0.01 (0.01, 0.02)	< 0.0001

Notes: $n = 62$. Deviance explained = 25%. Coefficients are expressed on the logit scale. *P* values and confidence intervals have been corrected for heteroskedasticity. Venues were weighted by the population-weighted number of respondents who visited that venue most frequently. There was interaction between the number of EGMs and venue type fitted in this model.

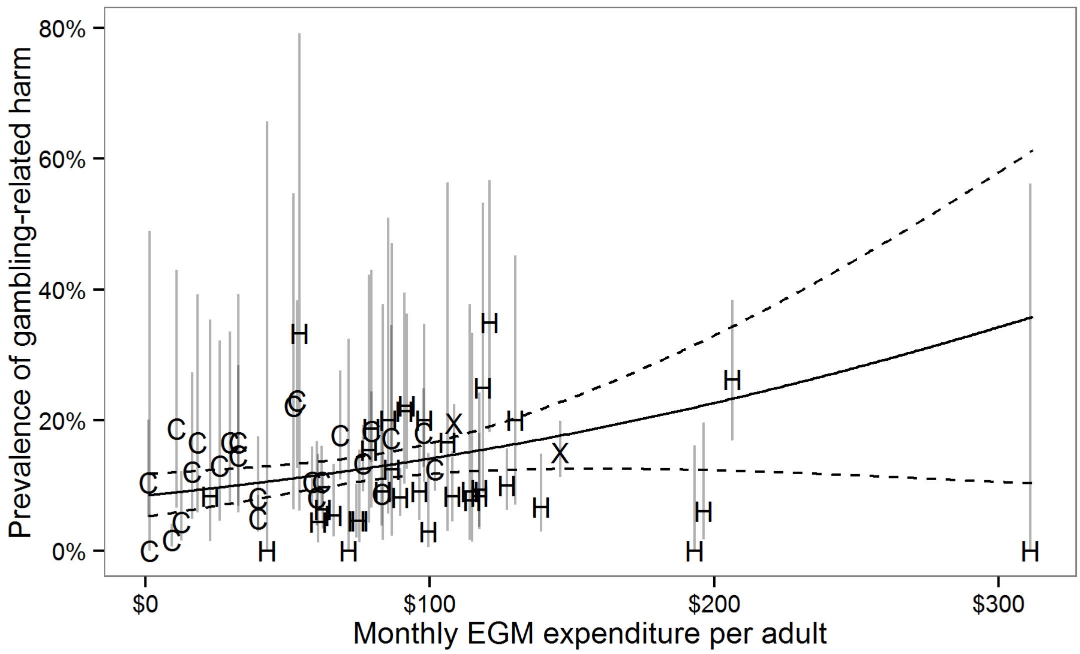

Figure 4.1: Predicted prevalence of gambling-related harm for a hypothetical club with the median number of EGMs (22). The solid black line shows the fitted regression line, and the dashed black lines outline the 95% confidence bounds. Points indicate actual venues in the study. Symbols X, C and H indicate venues of type casino, club and hotel, respectively. The intersecting vertical grey lines showing the 95% confidence interval for the prevalence of gambling-related harm at that venue, calculated using Wilson's method. Wilson's confidence intervals are asymmetric except when *P* = 0.5.

4.6 Discussion

The level of gambling-related harm varied substantially among venues, both between venues of different types (i.e. hotels, clubs and casinos) and within those categories. The prevalence of gambling-related harm at the venue level is significantly correlated with estimated monthly EGM expenditure per adult in both bivariate linear and multivariate binomial models. Holding all other variables constant, for a typical venue in our study area, each AUD20 increase in monthly EGM expenditure per adult is associated with an estimated 1.7% increase in the prevalence of gambling harm for a club with 22 EGMs.

These data are consistent with the hypothesis that EGM expenditure predicts the rate of gambling-related harm. While this is the first study of its kind and thus replication in other geographic contexts is needed, we cautiously suggest that the use of per capita EGM expenditure as a proxy for gambling-related harm may be justified. Furthermore, our findings are consistent with the prediction of the Total Consumption Theory, lending further support to its application in the domain of gambling.

We expect that the finding of a significant relationship between EGM expenditure and the prevalence of gambling-related harm at EGM venues is generalizable to other settings (and to other modes of gambling), wherever those experiencing gambling-related harm account for a substantial proportion of aggregate gambling expenditure. However, the precise magnitude of the relationship between expenditure and rates of harm is likely to vary between jurisdictions (and within the same jurisdiction over time) due to environmental, regulatory and social differences. Therefore, direct calculation of the proportion of EGM gamblers experiencing harm made from the coefficients estimated in this study should be undertaken with caution.

Although this cross-sectional study does not demonstrate a causal relationship between gambling expenditure and gambling-related harm, the correlation between EGM expenditure and gambling related-harm is important. We are not advancing a simplistic single-causal model in which visiting high expenditure venues *causes* disordered gambling pathology (although we do not rule out this possibility). Instead, we suggest that excessive gambling expenditure is conceptually and empirically inseparable from gambling-related harm because expenditure of money is the proximate source of many of the negative consequences associated with harmful gambling. Therefore, the money lost at EGM venues constitutes a harm in itself for some gamblers and this is detectable in aggregate gambling expenditure data.

4.6.1 Limitations

The relatively low response rate threatens internal validity in two ways. First, the sample composition is older, better educated and more likely to be female than the general population, meaning that the findings may be specific to this particular population subgroup. However, previous studies (Hansen and Rossow, 2008; Lund, 2008) and the Total Consumption Theory of gambling suggest that the relationship between gambling expenditure and gambling harm should be present in all population subgroups, even if harm rates vary among these groups. If this is the case, then the relationship between expenditure and harm should be robust to response bias. To investigate this proposition, we reanalysed our data on seven large subpopulations of respondents, and found little evidence to suggest the absence of a relationship between expenditure and harm in a population subgroup (see Appendix D, Figure D.1 and Table D.2). Therefore, we suggest that the substantive result of an association between expenditure and harm is not invalidated by this study's low response rate.

Second, the use of a mail survey and the recruitment method whereby any household member was eligible to reply to the questionnaire are all likely to skew the sample in favour of gamblers when compared to a telephone survey (Williams et al., 2012). This selection bias is likely to increase the estimated rates of gambling-related harm because gambling participation is the most important predictor of gambling-related harm. Indeed, our estimate of the rate of PGSI 8+ problem gambling in this study is several times that found in the last state wide prevalence telephone survey in the same jurisdiction (Young et al., 2008). As such, our coefficient estimates for the association between expenditure and harm rates may be biased upwards. Nevertheless, our finding of a strong positive relationship between expenditure and harm at the venue level is still likely to be valid unless selection bias affects venues differentially. This means that relative harm rates of gambling venues estimated on the basis of expenditure are unlikely to be affected by bias.

There are several other possible sources of non-sampling error. First, our measures of service populations are estimates only. Second, the populations served by venues are likely to differ non-randomly in terms of household income. It is reasonable to expect that lower income individuals will tend to experience gambling-related harms at lower levels of expenditure, thus biasing the magnitude of the estimated relationship downwards. The survey used in this chapter did not have an adequate measure of household income, making an analysis of how this relationship is mediated by income impossible. However, as Chapter 3 demonstrated using individual level data, the inclusion of household income does not remove the relationship between expenditure and harm. Third, although this study included a venue with an estimated monthly EGM expenditure per adult of over 300 AUD, 98% of respondents visited venues estimated expenditure of less than 150 AUD. Three of the four outlier venues are located in the extreme peri-urban fringe of Darwin, suggesting that gambling behaviour may differ in the peri-urban hinterlands or that the Huff model may be under-estimating the service-area populations of peri-urban venues. Consequently, shape of the expenditure/harm curve when expenditure levels are above 150 AUD is open to question. While exploratory modelling suggests that a slight lessening of the expenditure-harm relationship may exist above AUD 150 (see Appendix D, Figure D.2), further data collection is required to test this. Finally, visitors in non-residential accommodation are likely to be underrepresented in the study and may have different venue choice behaviour, decreasing the precision of parameter estimates.

4.6.2 Conclusions

Our finding of a measurable correlation between gambling-related harm and EGM expenditure, as predicted by Total Consumption Theory, has the potential to reduce the data collection required to research and regulate EGM gambling within a jurisdiction. These resources could usefully be redirected to other research or harm minimisation initiatives. If replication studies in other jurisdictions confirm our finding, we see little reason for those seeking to investigate the spatial patterning of gambling-related harm to continue to collect survey data on this topic. Rather, studies in this domain may reasonably rely on per capita gambling expenditure estimates and research effort currently employed to describe aggregate gambler behaviour could be redeployed in an effort to explain the patterns we see in gambling expenditure data. Further research is required to confirm this relationship, and to understand how it is mediated by other variables such as household income.

Chapter 5: A meta-regression analysis of 41 Australian problem gambling prevalence estimates and their relationship to total spending on electronic gaming machines

5.1 Foreword

This chapter presents the final of three studies examining the relationship between EGM gambling losses and problem gambling. This particular study investigates the relationship between EGM losses and problem gambling among Australian states and territories, using 41 separate state-level problem gambling prevalence estimates conducted over twenty-five years as the units of analysis. This scale of analysis is important because most gambling legislation in Australia operates at the state and territory level, resulting in considerable variation in per capita losses across jurisdictions.

This study follows on from those presented in Chapters 2 and 3. Further investigation of this relationship at the jurisdictional level is necessary because it is possible for EGM expenditure and harm to be correlated at the scale of the individual and the gambling venue but not for jurisdictions. This could occur if, for example, problem gambling is entirely the result of personality-related predispositions rather than the gambling environment. Consequently, this study tests whether the expenditure – harm relationship scales up to the level of the jurisdiction. The study found a moderate degree of evidence consistent with the existence of an EGM expenditure – harm association for states and territories. However, this result was complicated by two unexpected findings relating to (a) the degree of heterogeneity between problem gambling prevalence studies, and (b) the questionable validity of the 'moderate risk problem gambling' classification. Because of these unexpected findings which have implications for the conduct of population research into problem gambling, this study is written primarily for an epidemiological audience.

This study has been accepted for publication in *BMC Public Health* but had not been published at the time of submission. Its bibliographic details are:

Markham, F., Young, M., Doran, B. & Sugden, M. A meta-regression analysis of 41 Australian problem gambling prevalence estimates and their relationship to total spending on electronic gaming machines. *BMC Public Health*.

5.2 Abstract

Background

Many jurisdictions regularly conduct surveys to estimate the prevalence of problem gambling in their adult populations. However, the comparison of such estimates is problematic due to methodological variations between studies. Total consumption theory suggests that an association between mean electronic gaming machine (EGM) and casino gambling losses and problem gambling prevalence estimates may exist. If this is the case, then changes in EGM losses may be used as a proxy indicator for changes in problem gambling prevalence. To test for this association this study examines the relationship between aggregated losses on electronic gaming machines (EGMs) and problem gambling prevalence estimates for Australian states and territories between 1994 and 2016.

Methods

A Bayesian meta-regression analysis of 41 cross-sectional problem gambling prevalence estimates was undertaken using EGM gambling losses, year of survey and methodological variations as predictor variables. General population studies of adults in Australian states and territories published before 1 July 2016 were considered in scope. 41 studies were identified, with a total of 267,367 participants. Problem gambling prevalence, moderate-risk problem gambling prevalence, problem gambling screen, administration mode and frequency threshold were extracted from surveys. Administrative data on EGM and casino gambling loss data were extracted from government reports and expressed as the proportion of household disposable income lost.

Results

Money lost on EGMs is correlated with problem gambling prevalence. An increase of 1% of household disposable income lost on EGMs and in casinos was associated with problem gambling prevalence estimates that were 1.33 times higher [95% credible interval 1.04, 1.71]. There was no clear association between EGM losses and moderate-risk problem gambling prevalence estimates. Moderate-risk problem gambling prevalence estimates were not explained by the models ($I^2 \geq 0.97$; $R^2 \leq 0.01$).

Conclusions

The present study adds to the weight of evidence that EGM losses are associated with the prevalence of problem gambling. No patterns were evident among moderate-risk problem gambling prevalence estimates, suggesting that this measure is either subject to pronounced measurement error or lacks construct validity. The high degree of residual heterogeneity raises questions about the validity of comparing problem gambling prevalence estimates, even after adjusting for methodological variations between studies.

5.3 Background

5.3.1 Introduction and rationale

Total consumption theory, or single distribution theory as it is sometimes known, predicts that the incidence of gambling-related harm is related to the amount of time and money spent on gambling within a given jurisdiction (Lund, 2008; Rose and Day, 1990). This prediction derives from a postulate of total consumption theory that, at the population level, a fixed proportion of total gambling activity will result in harm. Any growth in gambling will accordingly produce a proportionate growth in harm. If this hypothesis is correct, it implies that gambling-related harm would be best prevented by reducing the gambling consumption of the entire population, not just those gambling to excess. A related implication is that changes in mean gambling consumption may be used as a proxy indicator for changes in problem gambling prevalence.

A small research literature has investigated the veracity of the propositions of total consumption theory as they relate to gambling. In a pioneering analysis of household expenditure surveys, Grun and McKeigue (2000) found that mean gambling losses in British geographic regions were strongly correlated with the proportion of the population losing an ‘excessive’ sum of money on gambling, both before and after the introduction of the National Lottery. Lund (2008), analysing three independent Norwegian samples, found similar correlations between average gambling frequency and the proportion of the population gambling very frequently.

These studies did not specifically examine gambling-related harms, a shortcoming addressed by both Hansen and Rossow (2008) and Markham *et al.* (2014c). Hansen and Rossow examined problem gambling among Norwegian adolescents grouped by school, and found a strong correlation among schools between average losses on slot machines and the reported prevalence of problem-gambling symptoms. Similarly, Markham *et al.* found that the reported prevalence of two or more problem-gambling symptoms among gamblers using electronic gaming machines (EGMs) in Australian venues was correlated with the average amount lost on EGMs in those venues.

Few studies, however, have systematically examined the relationship between gambling-related harm and gambling losses at the spatial scale of the regulatory jurisdiction (e.g. the country, state, territory, province, Bundesland, etc.). The jurisdictional spatial scale is important since jurisdictions comprise the territorial unit at which gambling is most frequently regulated, gambling losses are usually reported (e.g. Canadian Partnership for Responsible Gambling, 2015; Queensland Government Statistician's Office, 2016), and problem gambling surveys are usually conducted (Williams et al., 2012). In one notable example of a jurisdictional-level study, the Productivity Commission (1999) surveyed problem gambling prevalence in all Australian states and territories and compared these prevalence estimates to total non-lottery gambling losses in the same jurisdictions, finding a positive correlation. Unfortunately, this study was constrained by design to an examination of only eight prevalence estimates, limiting its generalizability.

Counter-examples to predictions of total consumption theory have been forwarded by Abbott (2006), who describes the reduction in problem gambling prevalence estimates in New Zealand over a nine year period during which total gambling losses increased substantially. Consequently, in a series of studies (Abbott, 2006, 2005; Abbott et al., 2016b, 1999), Abbott proposed an alternative hypothesis of 'adaptation', in which the prevalence of problem gambling tends to fall over time. The reasons for adaptation may include a decline in gambling participation as the novelty of a new gambling activity dwindles, decreased average duration of gambling problems through destigmatisation and improved treatment, changing cultural norms, increased knowledge of gambling-related harms, and the introduction of regulations such as in-venue smoking bans and caps on EGM numbers.

One explanation for the dearth of convincing evidence about the relationships between gambling harms and gambling losses at jurisdictional scales is the problem of inter-study heterogeneity, generally thought to result from a lack of methodological consistency between prevalence studies. Inter-jurisdictional comparisons may be compromised because problem gambling prevalence studies tend to use heterogeneous methods that limit comparability (Doughney, 2009; Markham and Young, 2016; Sassen et al., 2011; Shaffer et al., 1999; Williams et al., 2012). As Sassen *et al.* (2011) found in their systematic review of 39 studies, the decade between 2000 and 2010 saw little methodological convergence among prevalence studies. In practice, measurement differences between prevalence studies may be so great as to render comparisons between them invalid. Nevertheless, as an examination of almost any government-commissioned problem gambling prevalence study will demonstrate, comparisons between prevalence estimates are routinely drawn. Despite concerns regarding validity, almost every problem gambling prevalence study seeks to benchmark prevalence estimates against those within the same jurisdiction at a previous point in time, or within other jurisdictions at a similar point in time.

The problems inherent in comparing prevalence estimates have been recognised by some scholars, who have attempted to regularise these prevalence rates to account for methodological variations (Abbott et al., 2016b; Jackson et al., 2010; Stone et al., 2014; Williams et al., 2012). However, the validity of comparing regularised estimates has not yet been established because the amount of residual heterogeneity among studies *after* adjustment is unknown. This is important because problem gambling prevalence estimates are the primary means through which gambling-related harm is monitored by regulators and governments. If prevalence estimates cannot be meaningfully compared, this calls into question validity of the routine practice of monitoring problem gambling prevalence using surveys (Markham and Young, 2016).

If the total consumption theory of gambling is correct, then monitoring total gambling losses might provide an alternative means to track the changing incidence of gambling-related harm. If gambling losses present an accurate and precise proxy measure for problem gambling prevalence, then the necessity to routinely conduct problem gambling prevalence estimates to monitor population-level rates of harm might be reduced. Instead, population-level gambling losses could be monitored as a proxy indicator for the incidence of harm in the population.

5.3.2 Objectives

This study analyses the association between problem gambling prevalence estimates and gambling losses for Australian states and territories between 1994 and 2015. It aims to answer the following specific research questions:

1. Is there an association between EGM and casino gambling losses and problem gambling prevalence estimates in Australian states and territories?

2. What degree of heterogeneity remains in estimates of problem gambling prevalence after regularising for methodological variations, EGM gambling losses and year of survey?

5.4 Methods

A meta-regression approach was used to estimate the association between EGM and casino gambling losses and problem gambling prevalence estimates for Australian states and territories.

5.4.1 Setting

The units of analysis for this study were the eight states and territories of Australia. EGMs were introduced to these jurisdictions in a staggered manner, with New South Wales the first to legalise EGMs in 1956 (Australian Institute for Gambling Research, 1999). This was followed by a wave of legalisations, mostly in the 1990s, which left Western Australia as the only jurisdiction without EGMs by 1997. Other legal gambling commodities that are available in all jurisdictions include lotteries, casino table games, instant lotteries, scratch cards, and betting on races, sports and special events.

Australia was selected as a study site because its eight federal states and territories pioneered the routine conduct of problem gambling prevalence studies, alongside Canada and the United States (Williams et al., 2012). Consequently, there have been sufficient problem gambling prevalence studies conducted in Australia to warrant a meta-analysis of their results. The study was limited to Australian states and territories rather than including jurisdictions in multiple countries to minimise the potential differences among the populations surveyed.

5.4.2 Data

Two sets of data were required:

1. problem gambling prevalence estimates and the characteristics of the studies which produced these estimates, and

2. EGM gambling losses in the state or territory that temporally match each prevalence study.

Problem gambling prevalence studies for Australian states and territories were identified through a systematic review process. Prevalence studies are most frequently published as reports in the ‘grey literature’ rather than as peer-reviewed journal articles. Consequently, the search strategy primarily involved the identification of this relatively well-known corpus of problem gambling prevalence studies from previous inventories (Productivity Commission, 2010; Williams et al., 2012). The websites of Australian government bodies that have commissioned problem gambling prevalence studies were searched to identify further studies for examination, as were the reference lists of identified studies. The search revealed one study for which the full text was unavailable. The lead author of this study was contacted by email and conducted data extraction at the current authors’ request. To be eligible for inclusion, prevalence studies had to: 1) target the general population aged 18 years or older, 2) measure 12-month or 6-month problem gambling using a validated problem gambling screen, 3) report results for one or more whole states or territories in Australia, 4) have been published prior to 1 July 2016, and 5) report on independent samples rather than longitudinal studies measuring change among the same respondents over time. Data were extracted independently by two coders, MS and FM. In cases where these two coders disagreed, data were coded independently by a third coder MY and a consensus meeting held, as discussed by Orwin and Vevea (2009).

EGM gambling losses were selected as the predictor variable of interest in preference to total gambling losses because expenditure on different forms of gambling produces differing levels of harm. Previous research has shown that EGM and casino losses are more closely associated with problem gambling than either total gambling losses or losses on other gambling products (Markham et al., 2016; John W. Welte et al., 2004). Gambling losses on casino table games were included in the study as Australian player loss statistics for casinos do not distinguish between EGM and non-EGM gambling losses (Queensland Government Statistician’s Office, 2016)health . Gambling losses on EGMs during the year of survey fieldwork were extracted from *Australian Gambling Statistics, 32nd Edition*, a complete and authoritative administrative dataset

(Queensland Government Statistician's Office 2016). This dataset is compiled by Queensland Treasury, on the basis of aggregate tax records provided by each Australian state or territory government.

5.4.3 Measures

The following measures were extracted from each problem gambling prevalence study: a) the prevalence of 'problem gambling', where problem gambling was defined as Problem Gambling Severity Index (PGSI) ≥ 8 or South Oaks Gambling Screen (SOGS) ≥ 5, b) the prevalence of 'moderate-risk problem gambling', where moderate-risk problem gambling was defined as PGSI 3 – 7 or SOGS 3 – 4, c) the jurisdiction, d) the year during which data was collected, e) the administration mode of the survey, i.e. telephone or face-to-face, f) whether the SOGS (Lesieur & Blume 1987) or the PGSI (Ferris & Wynne 2001a) was used to assess problem gambling, g) the sample size of the survey, and h) any 'frequency threshold' used to select which respondents would be administered the problem gambling screen. A frequency threshold is a rule by which the problem gambling screening instrument is only administered to a subset of respondents, selected on the basis of their reported gambling frequency. For example, the screen may only be administered to those who gamble at least weekly, with less frequent gamblers being imputed a problem gambling score of zero. The prevalence of problem gambling and the prevalence of moderate-risk problem gambling were the key outcome variables of interest.

Problem gambling screen and administration mode variables were coded as dummy variables. The gambling frequency threshold variable was collapsed into four categories: weekly gambling, fortnightly gambling, monthly gambling and less than monthly gambling. The year the survey was conducted was subtracted from 2015 to calculate the age of the reported survey.

The measure of EGM gambling losses used was the sum of EGM gambling losses in hotels and clubs and all gambling losses in casinos, both expressed as a percentage of total household disposable income (HDI). This is an aggregate measure, with a single number reported for each state and territory in each year. It is calculated by expressing the total gambling expenditure on EGMs and in casinos for the jurisdiction as a percentage of HDI for the jurisdiction, where HDI is the total income accruing to the household sector less household sector taxes. HDI for each state and territory is reported annually in the *Australian System of National Accounts* (Australian Bureau of Statistics 2016). Because gambling losses were reported for fiscal years (spanning 1 July – 30 June) while survey dates were recorded for calendar years, losses for each calendar year were estimated by calculating the mean of losses for the two overlapping fiscal years.

5.4.4 Statistical analysis

Random effects meta-regression was used to estimate the partial correlation between problem gambling prevalence and EGM and casino losses, after adjusting for methodological variations in prevalence studies and the year of the survey. Meta-regression is an extension of meta-analysis, but has different aims. In general, the goal of a meta-analysis is to pool the varying results of primary studies and thereby arrive at a more accurate and precise estimate of a quantity of interest (e.g. the population prevalence of problem gambling). In contrast, a meta-regression analysis aims to understand what causes variation in the findings of primary studies, using procedures developed for regression analysis (Borenstein et al. 2009). In this context, we might interpret this study as contributing to an 'epidemiology of problem gambling prevalence studies' (Thompson & Higgins 2002).

A random effects model is a statistical extension to fixed effects meta-regression. While a fixed effects analysis relies on the assumption that the quantity of interest (e.g. the population prevalence of problem gambling) is truly consistent across all studies even if it is imperfectly measured, this is rarely the case. For example, the populations under study are unlikely to be identical in all relevant factors, even after covariates are adjusted for. A random effects specification is more conservative because it does not make this strong assumption. When applied, random effects meta-regression usually results in estimates with wider confidence intervals than fixed effects meta-regression (Borenstein et al. 2009).

The adopted statistical approach modelled the estimated prevalence of problem gambling and moderate-risk problem gambling in each study as a function of the following predictor variables: EGM gambling losses in the jurisdiction; problem gambling screen; administration mode; frequency threshold; and survey year. Variance inflation factors for predictor variables were all less than 4.0. A binomial model specification with a logistic link function was used. Following Higgins and Thompson and Borenstein and colleagues (Higgins & Thompson 2002; Borenstein et al. 2009), the random effects meta-regression model was specified as follows:

$$y_i \sim \text{Binomial}(p_i,\ ssize_i)$$

$$logit(p_i) = \alpha + \beta_k \cdot X_{ik} + \theta_i$$

$$\theta_i \sim \text{N}\left(0, \frac{1}{\tau^2}\right)$$

$$\tau \sim \text{Uniform}(0, 10)$$

$$w_i = \frac{y_i(\ ssize_i\ -\ y_i)}{ssize_i}$$

$$\sigma^2 = \frac{\sum_i w_i\ (n-1)}{\left(\sum_i w_i\right)^2 - \sum_i w_i^2}$$

$$I^2 = \frac{\tau^2}{\tau^2 + \sigma^2}$$

where: y_i is the number of problem gamblers identified in study i; $ssize_i$ is the sample size of study i; α is a constant intercept; X_{ik} is a matrix of k predictor variables and β_k is a commensurate vector of estimated regression coefficients including EGM gambling losses, study year and methodological variations; θ_i is a normally distributed random effect with a standard deviation between studies of τ (or equivalently a precision of $\frac{1}{\tau^2}$); n is the number of studies under analysis; and I^2 is the proportion of residual variation in the estimates of problem gambling prevalence that is due to heterogeneity between studies (rather than sampling variation). The uniform prior distribution ranging from 0.0 to 10.0 for τ was specified on the basis of the simulations carried out by Lambert and colleagues (Lambert et al. 2005). All models were implemented using 'gold standard' (Higgins & Thompson 2002) fully Bayesian estimation with *R* and *JAGS* (R Core Team 2015; Plummer 2015). JAGS code listings for all model specifications are available in Appendix G (Listings G.1 – G.6).

The estimates of β coefficients were modelled in three different ways in order to test the robustness of results to the provision of prior information about their values. In the first model, all β coefficients were estimated from the prevalence study data set in the usual manner of regression analysis, using 'weakly informative' priors distributions. This method has the advantage of minimising residual heterogeneity, but risks identifying spurious correlations (Higgins & Thompson 2004). In the second model, coefficients – except for the intercept, the EGM loss coefficient and the year of survey coefficient – were 'fixed' on the basis of estimates derived from the small literature concerned with their estimation. This is equivalent to the recent practice that has been applied to compare prevalence estimates between studies (e.g. Abbott, Stone, et al. 2016; Williams et al. 2012), where prevalence estimates are normalised by multiplying them by fixed adjustment factors. Fixing coefficients is advantageous as it forces control variables to be set at plausible values and reduces the effective degrees of freedom of the models. However, it admits no variance in coefficient estimates. Consequently, a third method which constitutes a compromise between the first two was also adopted, assigning 'informative' prior distributions to control variable parameter coefficients, regularising them within plausible ranges while still admitting uncertainty in their estimated values.

Prior distributions for β coefficients in all three models are listed in Table 5.1. Weakly informative priors were placed on the β coefficients for HDI loss and year of survey in all models. The standard deviations for the informative priors were inflated by a factor of four compared to those derived from meta-analysis, on the basis that the small number of studies synthesised in the meta-analyses were likely to lead to an over-estimation of the precision of these distributions. The unpublished meta-analyses that formed the basis of the informative prior distributions are attached as Appendix F. Each meta-regression model was estimated twice, first with the prevalence of problem gambling as the outcome variable, and second with the prevalence of moderate-risk problem gambling as the outcome variable.

These models address the two research objectives of this study. First, the player loss β coefficient can be interpreted as an indicator of the magnitude and direction of any association between aggregate EGM gambling losses and the prevalence of problem gambling. Second, the I^2 estimate describes the degree of heterogeneity that remains among problem gambling prevalence estimates *after* accounting for both sampling variability within individual prevalence studies and the predictor variables described

Table 5.1: Prior distributions placed on meta-regression parameter coefficients

	Models of problem gambling			Models of moderate-risk problem gambling		
	Weakly informative	Informative	Fixed	Weakly informative	Informative	Fixed
Intercept	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$
% HDI lost on EGMs and at casinos	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$
Years before 2015	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$	$N(0.00, 10^{12})$
Administered face-to-face	$N(0.00, 10^{12})$	$N(0.08, 1.60)$	$N(0.08, 10^{-48})$	$N(0.00, 10^{12})$	$N(0.73, 0.71)$	$N(0.73, 10^{-48})$
Used SOGS	$N(0.00, 10^{12})$	$N(0.48, 0.40)$	$N(0.48, 10^{-48})$	$N(0.00, 10^{12})$	$N(-0.53, 0.40)$	$N(-0.53, 10^{-48})$
Monthly frequency threshold	$N(0.00, 10^{12})$	$N(-0.03, 0.46)$	$N(-0.03, 10^{-48})$	$N(0.00, 10^{12})$	$N(-0.14, 0.23)$	$N(-0.14, 10^{-48})$
Fortnightly frequency threshold	$N(0.00, 10^{12})$	$N(-0.10, 0.40)$	$N(-0.10, 10^{-48})$	$N(0.00, 10^{12})$	$N(-0.38, 0.22)$	$N(-0.38, 10^{-48})$
Weekly frequency threshold	$N(0.00, 10^{12})$	$N(-0.40, 0.37)$	$N(-0.40, 10^{-48})$	$N(0.00, 10^{12})$	$N(-0.56, 0.24)$	$N(-0.56, 10^{-48})$

above. In other words, I^2 measures the inconstancy in prevalence estimates between studies rather than real variation in the prevalence of problem gambling. I^2 has a range of 0.0 to 1.0, with Higgins *et al.* (2003) suggesting that the values of 0.25, 0.5, and 0.75 indicate low, moderate and high degrees of heterogeneity, respectively.

5.5 Results

5.5.1 Selection

A total of 45 problem gambling prevalence estimates were identified, including eight state-level estimates derived the Productivity Commission's 1999 national survey (see Appendix E, Figure E.1 and Table E.1). Once ineligible studies were excluded, 41 studies were identified that estimated the prevalence of problem gambling, 40 of which also estimated the prevalence of moderate-risk problem gambling. All extracted data for each study are listed in Table 5.2.

Just over half of studies identified used the PGSI (n = 21), with the remainder using SOGS. The vast majority were administered by telephone (n = 37). The most commonly used gambling frequency thresholds for the administration of a problem gambling screen were a weekly threshold (n = 19) and the combined category of an annual threshold, a six-monthly threshold or no threshold at all (n = 19). Only two studies used a monthly threshold, while a single study used a fortnightly threshold. The median sample size was 4,303 (*Inter-quartile range (IQR)* = 1253 – 9408), with a total of 267,367 adults responding to the 41 surveys. The median survey year was 2001 (*IQR* = 1999 – 2008).

5.5.2 Outcome data and main results

The average non-regularised prevalence of problem gambling across all studies was 0.9% of adults (95% Credible Interval (Cr.I.) 0.8%, 1.1%), with the average non-regularised prevalence of moderate-risk problem gambling estimated to be 1.8% of adults (95% Cr.I. 1.5%, 2.1%) (see Appendix E, Figures E.2 and E.3). There was an extraordinarily large degree of heterogeneity among these studies, with I^2 for problem gambling and moderate-risk problem gambling estimated at 0.95 (95% Cr.I. 0.94, 0.95) and 0.94 (95% Cr.I. 0.93, 0.96), respectively.

Table 5.2: Problem gambling prevalence studies which met the eligibility criteria

State or territory	Year	Sample size	Prevalence of problem gambling (% of adults)	Prevalence of moderate-risk problem gambling (% of adults)	Administration mode	Screen	Gambling frequency threshold	Losses on EGMs and at casinos (% of HDI)
ACT	1999	708	2.06	2.54	Telephone	SOGS	Weekly	1.768
ACT	2001	5445	1.91	1.21	Telephone	SOGS	Weekly	1.587
ACT	2009	5500	0.50	1.50	Telephone	PGSI	Monthly	0.781
ACT	2014-15	7068	0.40	1.10	Telephone	PGSI	Six-monthly or less often	0.594
NSW	1995	1390	2.20	3.08	Doorknock	SOGS	Weekly	1.952
NSW	1997	1209	3.00	4.14	Doorknock	SOGS	Weekly	2.247
NSW	1999	2632	2.55	2.57	Telephone	SOGS	Weekly	2.726
NSW	2006	5029	0.80	1.60	Telephone	PGSI	Weekly	2.570
NSW	2008-09	9408	0.40	1.30	Telephone	PGSI	Six-monthly or less often	2.082
NSW	2011	10000	0.80	2.90	Telephone	PGSI	Six-monthly or less often	1.944
NT	1999	607	1.89	0.42	Telephone	SOGS	Weekly	2.009
NT	2005	5246	0.64	1.57	Telephone	PGSI	Weekly	2.129
QLD	1999	1518	1.88	4.13	Telephone	SOGS	Weekly	1.888
QLD	2001	13082	0.83	2.70	Telephone	PGSI	Six-monthly or less often	1.948
QLD	2003-04	30373	0.55	1.97	Telephone	PGSI	Six-monthly or less often	2.105
QLD	2006-07	30188	0.47	1.80	Telephone	PGSI	Six-monthly or less often	1.608
QLD	2008-09	14962	0.37	1.60	Telephone	PGSI	Six-monthly or less often	1.460
QLD	2011-12	15088	0.48	1.90	Telephone	PGSI	Six-monthly or less often	1.346
SA	1996	1206	1.24		Telephone	SOGS	Weekly	1.520
SA	1999	1013	2.45	0.57	Telephone	SOGS	Weekly	1.824
SA	2001	6045	1.88	1.36	Telephone	SOGS	Monthly	1.886

SA	2005	17140	0.40	1.20	Telephone	PGSI	Fortnightly	1.992
SA	2012	9508	0.60	2.50	Telephone	PGSI	Six-monthly or less often	1.370
TAS	1994	1220	0.82	1.97	Doorknock	SOGS	Weekly	0.811
TAS	1996	1211	2.89	5.70	Telephone	SOGS	Six-monthly or less often	0.934
TAS	1999	810	0.44	1.73	Telephone	SOGS	Weekly	1.609
TAS	2000	1223	0.90	1.55	Telephone	SOGS	Six-monthly or less often	1.743
TAS	2005	6048	0.73	1.02	Telephone	PGSI	Weekly	1.715
TAS	2007	4051	0.54	0.86	Telephone	PGSI	Weekly	1.438
TAS	2011	4303	0.70	1.80	Telephone	PGSI	Six-monthly or less often	1.189
TAS	2013	5000	0.50	1.80	Telephone	PGSI	Six-monthly or less often	1.052
VIC	1997	2000	1.00	1.30	Telephone	SOGS	Six-monthly or less often	2.489
VIC	1998	1737	1.50	1.10	Telephone	SOGS	Six-monthly or less often	2.707
VIC	1999a	1760	0.80	1.30	Telephone	SOGS	Six-monthly or less often	2.843
VIC	1999b	2227	2.14	2.65	Telephone	SOGS	Weekly	2.843
VIC	2003	8479	0.97	0.91	Telephone	PGSI	Weekly	2.598
VIC	2007	2012	1.40	2.80	Telephone	PGSI	Six-monthly or less often	2.175
VIC	2008	15000	0.70	2.36	Telephone	PGSI	Six-monthly or less often	2.083
VIC	2014	13554	0.81	2.79	Telephone	PGSI	All, annual or six-month	1.748
WA	1994	1253	0.56	0.48	Doorknock	SOGS	Weekly	1.285
WA	1999	1114	0.70	4.15	Telephone	SOGS	Weekly	0.758

Notes: SOGS = South Oaks Gambling Screen; PGSI = Problem Gambling Severity Index. Full bibliographic details for each study can be found in the Supporting Online Documentation as Table S1.

An association between the prevalence of problem gambling and EGM and casino gambling losses was apparent in the meta-regression model with weakly informative priors. Parameter coefficients, expressed as 'prevalence ratios' are displayed in Table 5.3. Prevalence ratios should be interpreted analogously to incidence rate ratios, and can be multiplied with the intercept to predict the value of the outcome variable for a given set of predictor variable values. Every increase of one per cent of household disposable income lost on EGMs and at casinos was associated with problem gambling prevalence estimates that were 1.35 (95% Cr.I. 1.04, 1.74) times higher. Placing informative priors on other meta-regression coefficients decreased the parameter estimate slightly to 1.33 (95% Cr.I. 1.04 1.71). Fixing coefficients related to methodological variations in prevalence studies slightly reduced the estimated association between prevalence and EGM and casino gambling losses and decreased the precision of the estimate, bringing the 'no association' prevalence ratio of 1.0 to within the 95% credible interval (1.29, 95% Cr.I. 0.98, 1.72). Posterior estimates of the associations between prevalence and losses are visualised in Figure 5.1.

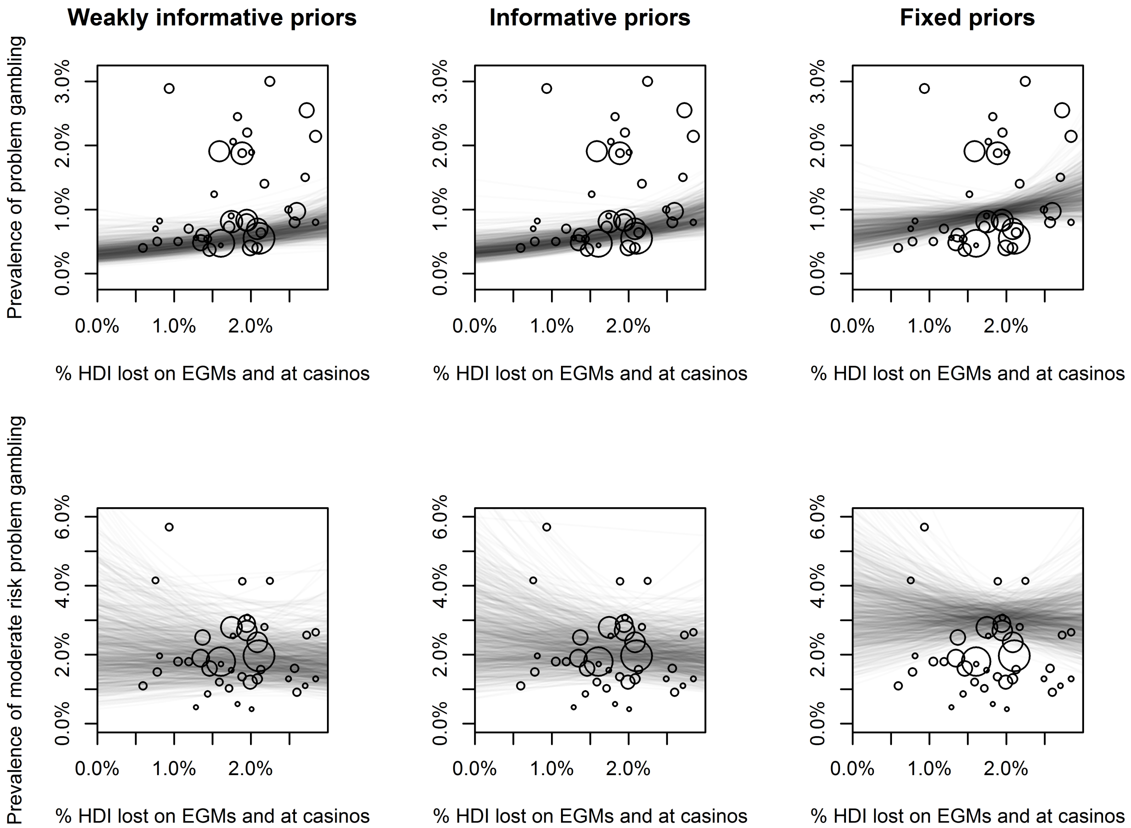

Figure 5.1: Posterior estimates of the association between prevalence and money lost gambling on EGMs and at casinos

Table 5.3: Meta-regression analyses of the prevalence of problem gambling and moderate-risk problem gambling. Parameter estimates have been exponentiated and should be interpreted as prevalence ratios, analogous to odds ratios.

Estimates of α and β coefficients and descriptive statistics	Problem gambling			Moderate-risk problem gambling		
	Weakly informative	Informative	Fixed	Weakly informative	Informative	Fixed
Intercept	**0.004 [0.002, 0.006]**	**0.004 [0.002, 0.006]**	**0.003 [0.002, 0.006]**	**0.021 [0.011, 0.042]**	**0.019 [0.010, 0.036]**	**0.016 [0.008, 0.032]**
% HDI lost on EGMs and at casinos	**1.35 [1.04, 1.74]**	**1.33 [1.04, 1.71]**	1.29 [0.98, 1.72]	0.94 [0.66, 1.34]	0.95 [0.68, 1.32]	0.97 [0.69, 1.37]
Years before 2015	0.99 [0.94, 1.04]	1.00 [0.96, 1.05]	**1.04 [1.02, 1.07]**	1.00 [0.93, 1.08]	1.02 [0.96, 1.08]	**1.05 [1.02, 1.09]**
Administered face-to-face	1.02 [0.59, 1.78]	1.02 [0.60, 1.74]	1.08 †	1.17 [0.55, 2.53]	1.33 [0.71, 2.55]	2.07 †
Used SOGS	**2.37 [1.37, 4.00]**	**2.18 [1.42, 3.34]**	1.62 †	1.20 [0.57, 2.52]	0.99 [0.55, 1.76]	0.59 †
Monthly frequency threshold	1.26 [0.68, 2.34]	1.15 [0.68, 1.84]	0.97 †	0.67 [0.28, 1.58]	0.81 [0.54, 1.20]	0.87 †
Fortnightly frequency threshold	0.65 [0.28, 1.50]	0.74 [0.43, 1.31]	0.91 †	0.65 [0.21, 2.15]	0.67 [0.45, 1.01]	0.68 †
Weekly frequency threshold	1.21 [0.86, 1.66]	1.10 [0.82, 1.47]	0.67 †	0.77 [0.49, 1.20]	**0.68 [0.50, 0.94]**	0.57 †
τ	0.37 [0.26, 0.51]	0.37 [0.25, 0.50]	0.46 [0.35, 0.60]	0.55 [0.40, 0.73]	0.53 [0.39, 0.70]	0.59 [0.44, 0.75]
R^2	0.65 [0.29, 0.89]	0.66 [0.31, 0.88]	0.47 [0.00, 0.71]	0.00 [0.00, 0.41]	0.00 [0.00, 0.44]	0.00 [0.00, 0.30]
I^2	0.87 [0.78, 0.93]	0.86 [0.78, 0.93]	0.91 [0.86, 0.95]	0.97 [0.96, 0.99]	0.97 [0.95, 0.99]	0.98 [0.96, 0.99]

Notes: 95% credible intervals are indicated in square brackets. Estimates where ,95% credible interval does not include 1.0 are indicated in boldface, except where coefficients are fixed. † indicates a coefficient that is fixed *a priori* rather than being estimated from the data. $n = 39$ for problem gambling and $n = 38$ for moderate-risk problem gambling. Problem gambling is defined as a PGSI score of 8 or more or a SOGS score of 5 or more. Moderate-risk problem gambling is defined as a PGSI score of 3 – 7 or a SOGS score of 3 – 4. HDI = household disposable income; EGM = Electronic Gaming Machine; SOGS = South Oaks Gambling Screen; PGSI = Problem Gambling Severity Index.

Few clear associations were found between moderate-risk problem gambling and any predictor variables using the models with weakly informative priors. With the use of informative priors, only the estimated association between moderate-risk problem gambling prevalence and a weekly gambling frequency threshold became reasonably precise (0.68, 95% Cr.I. 0.50, 0.94).

The number of years the study was conducted before 2015 was positively associated with both problem gambling prevalence and moderate-risk problem gambling prevalence, but only when meta-regression coefficients were fixed at the values arrived at from meta-analyses of previous within-study estimates. Models with fixed priors found that a one year increase in the age of the study was associated with prevalence estimates that were 1.04 (95% Cr.I. 1.02, 1.07) times greater for problem gambling and 1.05 (95% Cr.I. 1.02, 1.09) times greater for moderate-risk problem gambling.

A great deal of residual heterogeneity was evident among all models. While the models of problem gambling prevalence explained up to 66% of the variation among estimates, I^2 for these data still fell in the range between 0.78 and 0.95. This means that even after adjusting for covariates, a great deal of variation remained among problem gambling prevalence estimates, with unexplained heterogeneity between studies dominating random sampling error. This unsatisfactory situation was more extreme for estimates of moderate-risk problem gambling. No model explained even 1% of the variation among moderate-risk problem gambling prevalence estimates. The lower bound of estimates of I^2 for models of moderate-risk problem gambling was 0.95.

5.6 Discussion

The study had three key findings. First, problem gambling prevalence was associated with EGM and casino gambling losses in models with informative and weakly informative priors. An increase of 1% of household disposable income spend on EGMs and casino gambling associated with prevalence estimates that were approximately 1.3 times greater. In models where control parameter coefficients were fixed at values derived from meta-analyses, the point estimates of the gambling loss coefficient were similar but the 95% credible intervals widened to include a prevalence ratio of 1.0. In short, these results support the total consumption theory of gambling, but should be interpreted cautiously, given the degree of uncertainty evident in estimates.

The relatively wide uncertainty interval surrounding this finding is unsurprising given the relatively modest number of studies (*n* = 41) and their high degree of heterogeneity with respect to problem gambling prevalence ($I^2 > 0.85$ in all cases). Consequently, this study provides only a moderate degree of confidence that EGM gambling losses and problem gambling prevalence estimates are correlated. The best-fitting model, which used informative priors, suggested that an increase in EGM gambling losses of 1% of HDI is associated with a population-level increase in problem gambling prevalence of around 1.33 times (95% Cr.I. 1.04 – 1.71). The width of this uncertainty band is likely to be a consequence of measurement error overwhelming true variation in problem gambling prevalence. This high level of statistical noise in measurement is likely to derive from methodological variations that were unaccounted for in this study. For example, we did not adjust for way the survey was described to potential respondents, a factor that can impact non-response bias (Williams & Volberg 2009). Future studies using more consistently-collected data could seek to measure the relationship between EGM losses and problem gambling prevalence with more precision. These results support the need to phase out state-based prevalence studies and transition to national problem gambling prevalence studies that are adequately powered to investigate individual jurisdictions and that remain methodologically stable over time (Markham & Young 2016).

The moderate degree of uncertainty remaining around the association between EGM losses and harm should be interpreted in the context of parallel findings at other spatial scales. EGM gambling losses are correlated with risk of developing gambling problems for individuals (Markham et al. 2016) and for populations aggregated by county, school or gambling venue (Hansen & Rossow 2008; Markham, Young, et al. 2014). The present study adds to the weight of evidence that an increase in population losses on EGMs is associated with an increase in the prevalence of problem gambling.

The second key finding of the study was that a high degree of heterogeneity exists in problem gambling prevalence estimates. Only a moderate degree of variation among prevalence estimates was explained by EGM gambling losses, methodological variations, or year of study ($R^2 \leq 0.66$). When coefficient values for methodological variations were fixed based on prior research, variance explained fell substantially (R^2 = 0.47). Very little of the residual variation between problem gambling prevalence estimates was due to sampling error, with a very high degree of unexplained heterogeneity ($I^2 \geq 0.86$). Put differently, after adjusting for methodological variations, EGM losses and year of survey, no more than 14% of the residual differences between problem gambling prevalence estimates results from sampling error. These results raise questions about the validity of comparing problem gambling prevalence estimates, even after adjusting for methodological variations between studies.

The third key finding was the absence of any apparent pattern among moderate-risk problem gambling prevalence estimates. Contrary to expectations premised on total consumption theory, no link was evident between moderate-risk problem gambling prevalence estimates and EGM and casino gambling losses. Furthermore, no model explained any meaningful amount of the variation in moderate-risk problem gambling prevalence estimates, with the point estimate of R^2 falling below 0.01 in all cases. In other words, none of the models explained even 1% of the variation in estimates of moderate-risk problem gambling.

The lack of an apparent relationship between EGM gambling losses and moderate-risk problem gambling prevalence has two potential interpretations. It could be that there is no real relationship between moderate-risk problem gambling and EGM spending, in contradiction to total consumption theory. Alternatively, it is also plausible that the PGSI and SOGS are mismeasuring the population at moderate risk of problem gambling. Given that problem gambling screens have been developed and validated to identify problem or pathological gamblers (Ferris & Wynne 2001a), it may be that problem gambling screens are not fit for the purpose of identifying moderate-risk problem gamblers.

Several pieces of evidence offer tentative support for the mismeasurement conjecture. First, the only published validation study of the moderate-risk classification of the PGSI found that it lacked discriminant validity (Currie et al. 2013). In particular, this study found no practical differences between those scoring 1-2 on the PGSI and those scoring 3-7. Indeed, as McCready and Adlaf (2006) note, the PGSI does not include any items designed to discriminate among gamblers with less severe problems. As the authors of the PGSI acknowledge in their original study, the screen's division between low- and moderate-risk gambling categories is only tentatively supported by the survey data from which it was derived (Ferris & Wynne 2001a). In addition, the explanatory power of the variables included in the meta-regression analysis for predicting moderate-risk problem gambling prevalence estimates was exceptionally poor, with $R^2 < 0.01$ in all three models (Table 5.3). Similarly, residual heterogeneity was very high, with $I^2 \geq 0.97$ across all models. This implies that either moderate-risk problem gambling prevalence estimates are not impacted by methodological variations, or that such impacts are very small when compared to other unaccounted for factors. Finally, the meta-analysis of problem gambling screen effects presented in Appendix F finds a much greater degree of heterogeneity is present in estimates of screen impacts on moderate-risk problem gambling prevalence estimates ($I^2 = 0.91$) than problem gambling prevalence estimates ($I^2 = 0.69$). In other words, the 'within study' estimates of the impact of methodological variations on prevalence estimates vary a great deal *between* studies of moderate-risk problem gambling.

Taken together, this suggests that measures of moderate-risk problem gambling are extremely imprecise, to the point of possibly being the equivalent of statistical white noise. The apparent inability of current screening instruments to reliably identify this population is particularly problematic given that recent evidence suggests that this population experiences the greatest mass of gambling-related harms when measured in terms of disability-adjusted life years (Browne et al. 2016).

These results are subject to several limitations. First, a great deal of variation existed among prevalence estimates after adjusting for five predictor variables. Thus, estimates of any association between prevalence and EGM losses are necessarily imprecise. Second, the 41 prevalence estimates analysed in this study may be insufficient for the assessment of population trends, especially among the moderate-risk problem gambler population. Third, other methodological variations that this study was unable to adjust for may impact on problem gambling prevalence estimates. Finally, it is possible that some individuals may have been sampled in surveys in multiple years, especially in smaller jurisdictions. This is unlikely to have a substantial effect on the results.

5.7 Conclusions

This study has three key implications. First, the finding of an association between EGM and casino gambling losses and problem gambling prevalence is consistent with total consumption theory. Therefore, interventions by jurisdictional governments that reduce total EGM gambling losses among the whole population are likely to effectively reduce the prevalence of problem gambling. This result was evident despite the imprecision and heterogeneity of these estimates. Nevertheless, replication using a large cross-jurisdictional surveys that are consistent over time are required to confirm this association with a greater degree of confidence.

Second, this study demonstrates that using single-jurisdiction prevalence studies for making comparisons between jurisdictions or within the same jurisdiction over time is ineffectual (Doughney 2009). Even after deploying sophisticated statistical adjustments, the high degree of residual heterogeneity evident in this study suggests that the validity of comparing problem gambling prevalence estimates may be poor. The situation is far worse for moderate-risk problem gambling. It appears that although 267,367 Australian adults have responded to 41 surveys, we still are unable to confidently compare problem gambling prevalence either between jurisdictions or over time, a situation that is likely to be replicated internationally. It will be necessary to undertake adequately-powered, multi-jurisdictional prevalence studies – with survey instruments and data collection protocols that remain consistent over time – if the scientific value of problem gambling prevalence studies is to be increased in future.

Third, this study suggests that the suitability of PGSI and SOGS for estimating the population prevalence of moderate-risk problem gambling needs urgent investigation. The validity and reliability of the moderate-risk problem gambling classification is unclear. No patterns were evident among the estimates of the population prevalence of this group between studies, symptomatic of either extreme measurement errors or poor construct validity. The interpretation of moderate-risk problem gambling prevalence estimates should only be undertaken with extreme caution until a greater degree of conceptual and statistical clarity is brought to the identification of this population.

Chapter 6: Improving spatial microsimulation estimates of health outcomes by including geographic indicators of health behaviour: The example of problem gambling

6.1 Foreword

This chapter presents a method for estimating the prevalence of problem gambling in small areas, thereby realising the primary objective of this thesis. It builds on Chapter 2 directly, utilising the Huff model to estimate the distribution of EGM losses within each residential area. It then estimates the small-area prevalence of problem gambling on the basis of the Huff-model derived estimates, combining these with census and survey data. The validity of relying on EGM expenditure as a predictor of problem gambling prevalence was verified by the consistent relationship between expenditure and harm documented in Chapters 3 – 5.

The methods presented in this study constitute a substantial improvement over those that have previously been used to map problem gambling risk. For the first time, problem gambling risk has been estimated spatially on the basis of actual gambler behaviour and survey data, rather than assumed models of spatial behaviour and social vulnerability. Consequently, these methods are immediately transferrable for applied use in the context of EGM licensing and regulation.

However, in addition to its practical implications in the area of gambling regulation, this study has also contributed to refining spatial microsimulation methods themselves. This chapter focuses on methodological developments and is presented primarily for an audience of health geographers.

This study has been accepted for publication in *Health and Place* but had not been published at the time of submission. Its bibliographic details are:

Markham, F., Young, M. & Doran, B. Improving spatial microsimulation estimates of health outcomes by including geographic indicators of health behaviour: The example of problem gambling. *Health and Place.* doi: 10.1016/j.healthplace.2017.04.008

6.2 Abstract

Gambling is an important public health issue, with recent estimates ranking it as the third largest contributor of disability adjusted life years lost to ill-health. However, no studies to date have estimated the spatial distribution of gambling-related harm in small areas on the basis of surveys of problem gambling. This study extends spatial microsimulation approaches to include a spatially-referenced measure of health behaviour as a constraint variable in order to better estimate the spatial distribution of problem gambling. Specifically, this study allocates georeferenced electronic gaming machine expenditure data to small residential areas using a Huff model. This study demonstrates how the incorporation of auxiliary spatial data on health behaviors such as gambling expenditure can improve spatial microsimulation estimates of health outcomes like problem gambling.

6.3 Introduction

6.3.1 Background

Problem gambling, characterised by difficulties limiting time and money spent on gambling, is a significant and growing public health issue. Harms arising from problem gambling often include financial stress, deteriorated mental and physical health, strained interpersonal relationships, violence and crime. The serious nature of these impacts, combined with their relatively high prevalence in the population, means that problem gambling is in aggregate a serious public health burden. For example, problem gambling has been estimated to be the third-largest contributor to the burden of disability in Victoria, Australia, following major depression and alcohol abuse and dependence (Browne et al. 2016).

Despite its significance as a public health problem, little is currently known about the spatial distribution of problem gambling. Unpublished administrative data on gambling expenditure tends to show highly uneven spatial distributions, suggestive of gambling-related health inequalities. Yet few scholars have specifically examined the spatial distribution of gambling losses. One notable exception is Rintoul *et al.*'s (2013) study, which found that per capita electronic gaming machine (EGM) expenditure was highly concentrated in the most disadvantaged areas of Melbourne. More frequently, the spatial distribution of gambling venues has been mapped and correlated with indicators of deprivation or socioeconomic disadvantage. For example, studies of betting shops in London in 1966 (Newman 1972) and 2010 (Wardle et al. 2014) show that a historical spatial concentration in more deprived neighbourhoods continues to contemporary times. Similar spatial relationships between EGM venue density and disadvantage have been consistently observed in Australia, Canada, and New Zealand (e.g. Marshall & Baker 2002; Rush et al. 2007; Wheeler et al. 2006). Moreover, the relationship between venue density and disadvantage may be robust to changes in scale, with modest spatial correlations evident for small geographic zones (with an average of 225 dwellings), as well as for much larger spatial units with populations measured in the tens of thousands (Marshall & Baker 2001a).

The uneven provisioning of gambling venues and gambling expenditure suggests that the prevalence of problem gambling is also likely to be spatially patterned. Yet the degree to which the health burden of problem gambling is spatially uneven is currently unknown. Put simply, it is unclear if residents of some areas suffer from the adverse impacts of gambling more than others.

A spatial approach to modelling the prevalence of problem gambling is required in order to understand these geographic health inequalities. Beyond an academic imperative to understand the distribution of gambling harms, knowledge of the location of areas of high and low problem gambling prevalence would be useful for a range of practical applications. For example, licensing authorities are typically required to undertake local social impact assessments when new gambling venues are proposed, a task which is difficult to undertake in the absence of local data on the prevalence of problem gambling. Similarly, resources for treatment services ought to be provisioned on the basis of local needs. In short, there are both academic and practical imperatives to understand the spatial distribution of problem gambling.

Yet no studies to date have explicitly sought to estimate the prevalence of problem gambling in small areas. Five notable studies have, however, sought to map the distribution of what Welsh *et al.* term 'debtogenic landscapes' (2014) - urban environments conducive to, or symptomatic of, problem gambling. Taking a combinatory approach, Robitaille and Herjean (2008) mapped the demographic risk factors for problem gambling (i.e. gender, age, income, marital status, income, ethnicity and employment status) and found a spatial correlation between areas of high-risk demographics and the accessibility of gambling venues. Doran and Young (2010) undertook a conceptually similar study, but used index modelling and substituted an index of disadvantage derived using principle components analysis in place of Robitaille and Herjean's separate risk factor layers. This methodology that has since been replicated (Conway 2015). Rintoul *et al.* (2013) extended this approach, weighting accessibility scores for venues by the volume of EGM expenditure within those venues, rather than following Doran and Young's approach of weighting venues by number of EGMs. The most comprehensive study to date has been that of Wardle *et al.* (2016). This study produced a weighted linear combination of a wide range of risk factors for, and indicators of, problem gambling. They measured not just socio-demographic risk but also the location and utilisation of various mental health services (including problem gambling treatment), the residential location of people utilising homelessness services, and the location of payday-loan outlets and food banks.

The strength of these studies is that they capture the spatial variations of a wide range of gambling-related variables. However, their chief shortcoming is that they are entirely predictive. The outcome variable they produce is a unitless measure of vulnerability, but this index is not calibrated against any empirical data on outcomes *per se*. Consequently, the weights that are assigned to the various elements of vulnerability indices are necessarily arbitrary, with no empirical grounding beyond expert opinion. In effect, they operate in a manner similar to a spatial version of multiple linear regression in which all coefficient values are determined *a priori* by the analyst rather than being estimated from data. At best, the maps produced using this approach provide an educated guess regarding the location and relative prevalence of problem gambling.

This shortcoming is unfortunate given the collection of a large quantity of survey data specifically designed to investigate problem gambling (Williams et al. 2012). The primary limitation of existing surveys that hinders their use in the production of small-area estimates of problem gambling is that they are typically not geocoded (or geocodes are obscured for privacy reasons), so it is difficult to precisely allocate survey responses to residential locations. Even where geocodes are provided, surveys generally do not collect sufficiently spatially-dense data to produce estimates of harm at fine spatial resolutions using regression-based methods such as multilevel modelling (Whitworth et al. 2016).

Other methods such as spatial microsimulation provide an attractive means of producing small area estimates. This paper shows how the strengths of the index modelling approaches discussed above can be combined with well-developed spatial methods to improve small-area estimates. Specifically, spatial microsimulation is used to produce empirically-calibrated small-area estimates of problem gambling that take advantage of spatially-referenced administrative data as well as census data to constrain estimates.

6.3.2 Improving spatial microsimulation estimates of health outcomes with geographic indicators of risk

Spatial microsimulation provides a suite of methods for geographically allocating survey responses to small spatial areas using well-defined spatial data about the small areas to constrain estimates. The purpose is to synthesise a set of geographically-specific study populations, which can then be further analysed in a manner relevant to the study domain and research questions (Lovelace & Dumont 2016). In typical usage, spatial microsimulation involves three discrete steps. First, the total counts of persons across different socio-demographic categories are extracted from a population census at the finest possible geographic scale, either as counts of a single census category or as counts from a cross-tabulation of two or more variables. Second, these census-derived totals are harmonised with variables measuring the same construct (e.g. sex, age bracket, etc.) from a survey for which unit record data are available. The outcome variables of interest, which are measured by the survey but not the census, are also identified and included in the unit record data. Third, spatial microsimulation methods are used to allocate survey responses to small areas in a manner that makes the synthesised small-area totals match the census margins as closely as possible. This enables reliable estimates of the outcome variables of interest to be produced at finer geographic scales than those possible using the survey alone.

Spatial microsimulation has been used in this manner to produce small-area estimates of a range of health outcomes. For example, Cataife (2014) combined survey data with census statistics to produce estimates of the prevalence of obesity in tracts spanning just a few city blocks. Similarly, Smith *et al.* (2011) estimated smoking prevalence in Census Area Units in New Zealand, synthesising a national health survey with census data on four socio-demographic variables. These examples share a standard approach to spatial microsimulation in which survey responses are combined with census data without recourse to other sources of spatial information.

However, the reliability of the estimates produced by these methods depends in large part on the ability of census variables to predict the health outcome of interest. In general, the choice of constraint variables is crucial in producing reliable spatial microsimulation based estimates (Smith et al. 2011). In cases where the outcome measure is strongly related to a small number of census variables or their interactions, spatial microsimulation is likely to produce good results. However, for many policy-relevant problems, the outcome of interest is only poorly correlated with census variables. This makes the use of spatial microsimulation less attractive and suggests a need for further, spatially-referenced constraint variables that may not be provided in population censuses.

Environmental risk factors play a role in mediating many health outcomes and provide a likely candidate for providing such additional information. Variables measuring environmental risk factors (e.g. EGM accessibility) have been incorporated into the problem gambling index models described above (e.g. Conway 2015; Doran & Young 2010; Robitaille & Herjean 2008). In a study aimed at estimating the uptake of gestational diabetes screening in small areas in Ireland, Cullinan *et al.* (2012) provide an example of how auxiliary spatial information on risk can be used to augment typical spatial microsimulation approaches. Because screening uptake is highly dependent on the spatial accessibility of screening facilities, an application of spatial microsimulation to census data alone would have provided geographically questionable results. Therefore, using geocoded hospital register data, the authors converted absolute spatial measures (i.e. individuals' residential latitude and longitude) into a relative spatial measure (i.e. distance to nearest screening centre) and incorporated this as a constraint variable into their model. They were also able to extract other contextual variables such as urban or rural status for each person in the register on the basis of their residential location. These relative-spatial attributes from the register were combined with census data and GIS-calculated data using spatial microsimulation to produce improved small area estimates of screening rates. As this study demonstrated, the inclusion of spatial information above-and-beyond census marginal totals is possible and may indeed be required in some cases to generate sensible spatial microsimulation models.

However, the best predictors of health outcomes are often health-related behaviours. For example, in the case of gambling, socio-demographic variables typically explain around 10% of variance in the problem gambling classification of individuals, while the inclusion of gambling expenditure variables increases variance explained to around 30% (Markham et al. 2016). Geographic indicators of health behaviours are sometimes available and have been included in spatial index models of vulnerability (e.g. Rintoul et al. 2013; Wardle et al. 2016), but rarely in spatial microsimulation studies. We suggest that the inclusion of health behavioural variables in spatial microsimulation analyses is likely to improve the reliability of small area estimates. If surveys provide measures of health behaviours as well as health outcomes, then spatial data relating to health behaviours can provide a crucial link between aggregate collective behaviour at the small area and aggregate health outcomes. This requires the creation of constraint variables for small areas measuring health behaviours that can augment census-derived marginal totals. Census-derived and health-behavioural constraints can then be combined with survey data in spatial microsimulation models.

This solution poses additional problems at the data processing stage. In particular, the transformation of aggregate data relating to health behaviours into categorical constraints for small areas is not always straightforward. We suggest that the answer to these questions is likely to be domain specific. In the case of gambling, the spatial behaviour of consumers is already reasonably well understood (Markham, Doran, et al. 2014a), meaning that point-based data on EGM expenditure can be converted to mean per capita expenditure estimates for small areas on the basis of a statistical model. The conversion of population means to numbers of people in different gambling involvement categories can be made on the basis of the distribution of behavioural measures in the survey itself. This prior knowledge can be used as the basis for estimates of mean gambling losses in small areas. Analogous, domain-specific conversations are likely to be possible for other research problems.

6.3.3 Objectives

This study aims to demonstrate the potential for geographical indicators of health behaviours to improve small area estimates derived using spatial microsimulation, with reference to the particular example of estimating problem gambling prevalence. Specifically, this study aims to:

1. compare the explanatory power of individual level models with models including the following predictor variables: a) census variables, b) environmental risk factors, c) health-behavioural measures, and d) a combination of the most important variables across the three categories.

2. compare small-area estimates produced using spatial microsimulation across these three model configurations.

These objectives are pursued in the context of estimating the prevalence of problem gambling in small census areas.

6.4 Materials and methods

6.4.1 Setting

This setting for this study is the urban areas of the Northern Territory (NT) of Australia, primarily the towns of Darwin, Katherine and Alice Springs, and their peri-urban hinterlands. At the time of data collection, 88% of EGMs in the NT were located in or adjacent to these three towns, dispersed across 64 licensed gambling venues. The two largest EGM venues in the study area were the casinos in Alice Springs and Darwin, which together contained more than half of the approximately 2000 EGMs in these towns. The remaining EGMs were distributed among 36 hotels (with approximately 350 EGMs) and 26 clubs (with over 600 EGMs). Clubs are formally not-for-profit community centres, such as sporting or returned servicepersons clubs and were allowed a maximum of 45 EGMs per venue. Hotels or pubs are commercial businesses and were limited to a maximum of 10 EGMs per venue. The EGMs offered by these venues – known as ‘poker machines’ in the Australian vernacular – were high-intensity slot machines, with no minimum spin rate and a maximum bet of $5 per spin, resulting in an average cost of high-intensity gambling of approximately $600 per hour (Productivity Commission 2010). EGMs can be loaded with up to $1000 at a time. No regulations enforced limit setting by gamblers or breaks in gambling sessions.

6.4.2 Data

The primary data set of interest is a geocoded survey conducted in the urban areas of the NT. Between April and September 2010, a questionnaire was mailed to all 46,263 households in the study area to which Australia Post would deliver unsolicited mail, and a further 2300 questionnaires were hand delivered to peri-urban addresses beyond the range of the postal service. The sample frame was derived from the Australian geocoded national address file (G-NAF), and excluded areas zoned for non-residential uses. Any adult in the household was eligible to participate. The Human Research Ethics

Committee of Charles Darwin University granted approval to conduct the study (protocol no. H09048). The questionnaire elicited information on socio-demographics (age, sex, Indigenous status, marital status, and education), gambling behaviour (venues visited, and EGM gambling participation, frequency and session length), and problem gambling (measured using the Problem Gambling Severity Index or PGSI: Ferris & Wynne 2001b). Because the G-NAF was used as a sample frame, all responses could be precisely geocoded to the dwelling level with a 100% match rate. Neighbourhood disadvantage was measured on the basis of residential location, using a census-derived index of economic resources (IER). The IER is produced using a principal components analysis by the Australian Bureau of Statistics (2013) at the Statistical Area 1 (SA1) level of aggregation, a spatial unit with a median population of approximately 400 people. IER values in the study area were discretised into terciles, with the lowest tercile representing areas with the fewest economic resources.

Data to match the survey questions on age, sex, Indigenous status, marital status and education were derived from the 2011 Australian Bureau of Statistics Census of Population and Housing at the SA1 level of aggregation. Age and sex were cross-tabulated to produce separate marginal totals for age brackets (18-39 years, 40-54 years and 55 years or older) for each sex. All other variables were extracted as total counts of single variables for each SA1.

Accessibility to EGM venues is a well-documented environmental risk factor of problem gambling (e.g. Pearce et al. 2008; Welte, Wieczorek, et al. 2004; Young et al. 2012b). An EGM accessibility surface was developed using an unconstrained spatial interaction model by maximising the log-likelihood equation derived by Fotheringham and O'Kelly (1989). EGM venue locations were manually geocoded and a range of attractiveness variables were collected, including: number of EGMs, venue license category, whether the venue was a tourist-oriented inner city bar, proximity to shopping centres, distance to the central business district, and whether the venue had ocean views. Using participants' responses to a question about which EGM venues they visited in the last 30 days, model parameters were estimated that best predicted their reported travel behaviour. The calibrated accessibility model is presented in Equation 1, where: $access_j$ indicates the accessibility score of respondent j; d_{ij} indicates the distance in km between

EGM venue i and the home of respondent j; and other indicators represent the attractiveness variables described above.

$$access_j = \sum_i d_{ij}^{-0.83} \cdot size_i^{0.89} \cdot club_i^{0.48} \cdot casino_i^{-0.06} \cdot touristbar_i^{-0.25} \cdot shopcentre_i^{0.20} \cdot \log(dist\_cbd)_i^{0.24} \cdot ocean_i^{0.23} \quad (1)$$

The calibrated parameters of the accessibility model indicate that propensity to visit venues is only weakly related to distance to venue. The gradient of the distance decay curve is relatively flat, with an exponent of -0.83 (95% C.I.: -0.81, -0.85) suggesting that accessibility impacts on visitation behaviour at a regional scale rather than a highly localised scale (Hansen 1959). Venue size is also crucial to accessibility, with a venue with 45 EGMs contributing 3.8 (95% C.I.: 3.5, 4.2) times more to accessibility than a venue with 10 EGMs. Clubs also contributed more to EGM accessibility than hotels, while venues that were located close to supermarkets or that had ocean views were also more accessible. Accessibility scores were calculated for each respondent and discretised into terciles.

The health behaviour of interest was gambling involvement. Involvement was measured for local areas using per capita gambling expenditure. Mean per capita gambling expenditure in each SA1 was estimated from administrative data on EGM expenditure for individual venues during the survey period provided by the NT Department of Justice. These authoritative data are considered complete and reliable because they are generated from a computerised centralised monitoring system. The monitoring system collects real-time transaction data from each EGM in the NT, an arrangement designed to prevent EGMs being used to facilitate organised crime (Australian Institute for Gambling Research 1999). A previously published Huff model that was calibrated on the same data set was used to allocate expenditure at EGM venues to local areas (Markham, Doran, et al. 2014a; Markham, Young, et al. 2014), producing an estimate of mean per capita expenditure for each SA1. The mean expenditure estimates were used as a basis for calculating the number of persons with differing gambling involvement levels in each SA1. Specifically, minutes spent gambling on EGMs in the last 30 days was calculated for each respondent on the basis of their survey responses. Expenditure was derived via time because the survey instrument did not contain questions about money lost gambling. Time spent gambling was converted to dollars spent gambling using the population-weighted mean EGM expenditure velocity, which was calculated to be \$2.35 per minute. Survey estimated dollars spent per month for individuals and mean per capita EGM expenditure were combined at the SA1 level. A bivariate regression analysis was conducted on these SA1-level data to estimate the relationship between mean per capita expenditure and the percentage of residents with high EGM gambling involvement (defined as expenditure of \$300 in the last 30 days) or no EGM gambling involvement in the last 30 days. The remaining individuals spending \$1-\$299 in the last thirty days were classified as low involvement (see Figure 6.1), and was calculated as the remaining population in each SA1 after the other two groups had been accounted for. These relationships were used to estimate the number of people in each gambling involvement category in each SA1.

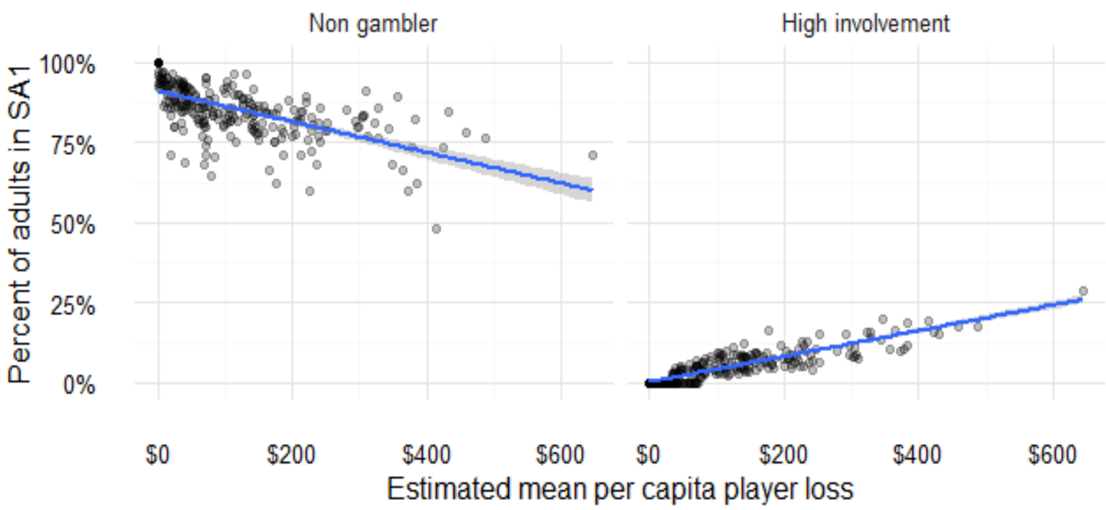

Figure 6.1: Estimated relationship between Huff model derived mean per capita EGM expenditure and percentage of respondents with survey-derived gambling involvement. Units of analysis were SA1s. Regression lines are weighted by the number of survey respondents in each SA1. R^2 for non-gamblers and high involvement categories were 0.41 and 0.78, respectively.

Ultimately, two data sets were assembled, each covering the same data items at different scales of aggregation, one for individuals primarily derived from the survey, and another of total counts aggregated to the SA1 level. A summary of these variables and their data sources is provided in Table 6.1.

Table 6.1: Summary of variables used in the spatial microsimulation analysis and their data sources

Variable	Individual-level source	SA1-level source
Age and sex	Survey	Census
Indigenous status	Survey	Census
Education	Survey	Census
Marital status	Survey	Census
Neighbourhood disadvantage	ABS IER for SA1 of survey respondent's residence	ABS IER, discretised into terciles.
Accessibility	Calculated for each survey respondent and discretised into terciles.	Calculated for each dwelling in the sample frame, with the proportion of dwellings in each SA1 in each discrete category calculated, with the number of persons in each category imputed from these proportions.
EGM gambling involvement	Survey	Calculated from a Huff model which allocates administrative data on expenditure in EGM venues to SA1s. SA1 mean per capita expenditure is converted into numbers of persons in each involvement category using the regression estimates presented in Figure 6.1. Categories were "No EGM participation", "Low EGM participation" and "High EGM participation".
Problem gambling	Survey	Not applicable, this is the outcome variable.

6.4.3 Statistical analysis

A two-phase approach was taken to the statistical analysis. In the first phase, the power of three sets of variables (socio-demographic, environmental and involvement) to explain problem gambling risk was explored. All models were limited to four predictor variables as previous research has suggested that the inclusion of too many constraints reduces the performance of spatial microsimulations (Tanton & Edwards 2012). Four logistic regression models were fit to predict whether or not an individual would meet the conventional classification of a problem gambler (a score eight or more on the PGSI). The first model contained socio-demographic variables only. The second model contained only two environmental risk factors, accessibility and neighbourhood

Table 6.2: Multiple logistic regression coefficients and indices of model fit for four different sets of variables predicting problem gambling among individuals

	Socio-demographic model		Environmental risk factor model		Gambling involvement model		Combined model	
	O.R.	95% C.I.	O.R.	95% C.I.	O.R.	95% C.I.	O.R.	95% C.I.
Intercept	0.02	0.01, 0.04	0.03	0.02, 0.04	0.01	0.01, 0.01	0.01	0.00, 0.01
Female, aged 18-39 years	1.00						1.00	
Female, aged 40-54 years	1.69	0.93, 3.19					1.47	0.78, 2.87
Female, aged 55 years or older	1.15	0.58, 2.29					0.66	0.33, 1.35
Male, aged 18-39 years	**6.29**	**3.46, 11.90**					**5.31**	**2.78, 10.51**
Male, aged 40-54 years	1.95	0.98, 3.93					1.60	0.77, 3.33
Male, aged 55 years or older	1.43	0.73, 2.84					0.84	0.42, 1.73
Married or in a de facto marriage	**0.51**	**0.36, 0.72**					**0.55**	**0.38, 0.80**
Indigenous	**4.20**	**2.46, 6.89**					**3.13**	**1.70, 5.54**
Attained school-level qualifications	1.00							
Attained technical qualifications	**0.48**	**0.28, 0.80**						
Attained university qualifications	**0.53**	**0.36, 0.77**						
Low accessibility tercile			1.00					
Medium accessibility tercile			1.23	0.81, 1.90				
High accessibility tercile			1.39	0.93, 2.11				
Low I.E.R. tercile			1.00					
Medium I.E.R. tercile			**0.62**	**0.43, 0.90**				
High I.E.R. tercile			**0.42**	**0.27, 0.64**				
Non E.G.M. gambler					1.00		1.00	
Spent between $1 and $300 on E.G.M.s in last 30 days					**6.43**	**3.95, 10.35**	**6.14**	**3.62, 10.20**
Spent $300 or more on E.G.M.s in last 30 days					**40.55**	**27.46, 60.66**	**43.52**	**28.42, 67.60**
A.I.C.	1260		1422		1111		985	
Pseudo R^2	0.14		0.02		0.23		0.33	

Notes: O.R. = odds ratio, C.I. = confidence interval, I.E.R. = index of economic resources, E.G.M. = electronic gaming machine, A.I.C. = Akaike's Information Criterion. All reported odds ratios are adjusted for other variables in the model. Bold type indicates odds ratios whose 95% confidence intervals do not contain 1.0.

disadvantage. The third model included only a single measure of health behaviour, 'gambling involvement', defined on the basis of EGM expenditure. The final model included the four predictor variables across all categories that best fit the data drawn from a total of seven possible predictor variables. Akaike's Information Criterion (AIC) and McFadden's pseudo-R^2 were reported as relative measures of model fit. Multicollinearity among predictor variables was unusually low, with generalised variance inflation factors calculated as less than 1.5 in all cases.

Finally, four spatial microsimulations were run using the same four combinations of predictor variables as the logistic regression analysis. Combinatorial optimisation was used to allocate individual survey respondents to SA1s. An implementation of simulated annealing was used to minimise total absolute error when synthesising the population of each SA1, with the maximum number of iterations set to 1000. The *sms* package in *R* was used to undertake this computation (Kavroudakis 2015). The prevalence of problem gambling was calculated among these synthesised populations at the SA1 level. SA1-level estimates from the four models were compared cartographically and formally tested for statistical correlations. Combinatorial optimisation was selected *a priori* as the method to undertake the microsimulation analysis rather than a generalised regression approach in order to reduce the convergence problems sometimes associated with the latter method. While combinatorial optimisation has the disadvantage of being stochastic, this was not seen as a serious drawback given the use-case for these estimates did not require perfect replicability.

6.5 Results

In total, 7049 people completed the survey, resulting in a response rate of 14.5%. The prevalence of problem gambling in the entire sample was 2.1% (95% C.I. 1.8%, 2.5%). Respondents were more likely to be female (61.9%, 95% C.I. 60.7%, 63.0%), and aged 55 or older (36.6%, 95% C.I 35.5%, 37.8%) than the general population in the study area. Mean imputed 30-day EGM gambling expenditure, calculated from reported time spent gambling, was \$114 among the sample (SD=\$470). The distribution of expenditure was highly right skewed as is typical for this kind of data, with 86.7% of the sample not participating in EGM gambling at all.

The logistic regression models of problem gambling show that gambling involvement is the single best predictor of problem gambling among individuals (see Table 6.2). The model including only gambling involvement explained 23% of the variance in problem gambling classification. In contrast, socio-demographic variables and environmental risk variables explained just 14% and 2% of variance, respectively. Combining the gambling involvement variable with selected socio-demographic variables produced the best fitting model, which explained 33% of the variance in problem gambling classifications. The four variables to include in this combined model were selected on the basis of AIC.

All variables except accessibility were significantly correlated with problem gambling risk. In particular, men aged between 18 and 39 were 5–6 times more likely to be problem gamblers than women of that same age. Indigenous people were 3 or 4 times more likely to report problem gambling. Those who were married or in facto relationships, those who had completed post-school education, and those who lived in wealthier areas were half as likely to report problem gambling. Finally, those who were imputed to have spent \$300 or more on EGM gambling in the last 30 days were 40 times more likely to report problem gambling than those who didn't gamble on EGMs in the same period.

The problem gambling prevalence estimated by the four spatial microsimulation were rather similar when analysed collectively for the entire study area. The socio-demographic model estimated problem gambling prevalence at 2.3%, the environmental risk factor model estimated 2.0%, the gambling involvement model estimated a prevalence of 2.1%, while the combined model estimated a population problem gambling prevalence of 2.2%. However, the spatial patterning of problem gambling changed substantially depending on model configuration. As Table 6.3 demonstrates, the prevalence of problem gambling at the SA1 level was only significantly correlated for the combined model and the socio-demographic model. Even in this case, the correlations were weak. This divergence is evident when problem gambling prevalence estimates are mapped spatially. Figure 6.2 shows the spatial distribution of problem gambling prevalence in north Darwin, an important region in the study area. In some areas, model predictions are relatively consistent, for example southern Tiwi, where prevalence estimates ranged from between 1.6% (Panel A) and 2.0% (Panels D). In contrast, prevalence estimates varied substantially between models in other areas like part of northern Leanyer, with estimates ranging from 0.7% (Panel B) to 4.1% (Panel A). In general, both Table 6.3 and Figure 6.2 demonstrate that the environmental risk factor model produces results that are dissimilar to those produced by the other three models.

Table 6.3: Correlation matrix of problem gambling prevalence estimates for SA1s produced using four spatial microsimulation models

	Socio-demographic model	Environmental risk factor model	Gambling involvement model	Combined model
Socio-demographic model	1.0			
Environmental risk factor model	-0.02	1.0		
Gambling involvement model	-0.02	0.03	1.0	
Combined model	**0.25**	-0.10	-0.01	1.0

Notes: Pearson's correlation coefficients are reported. Bold type indicates correlations that are significant at the $p < 0.05$ level.

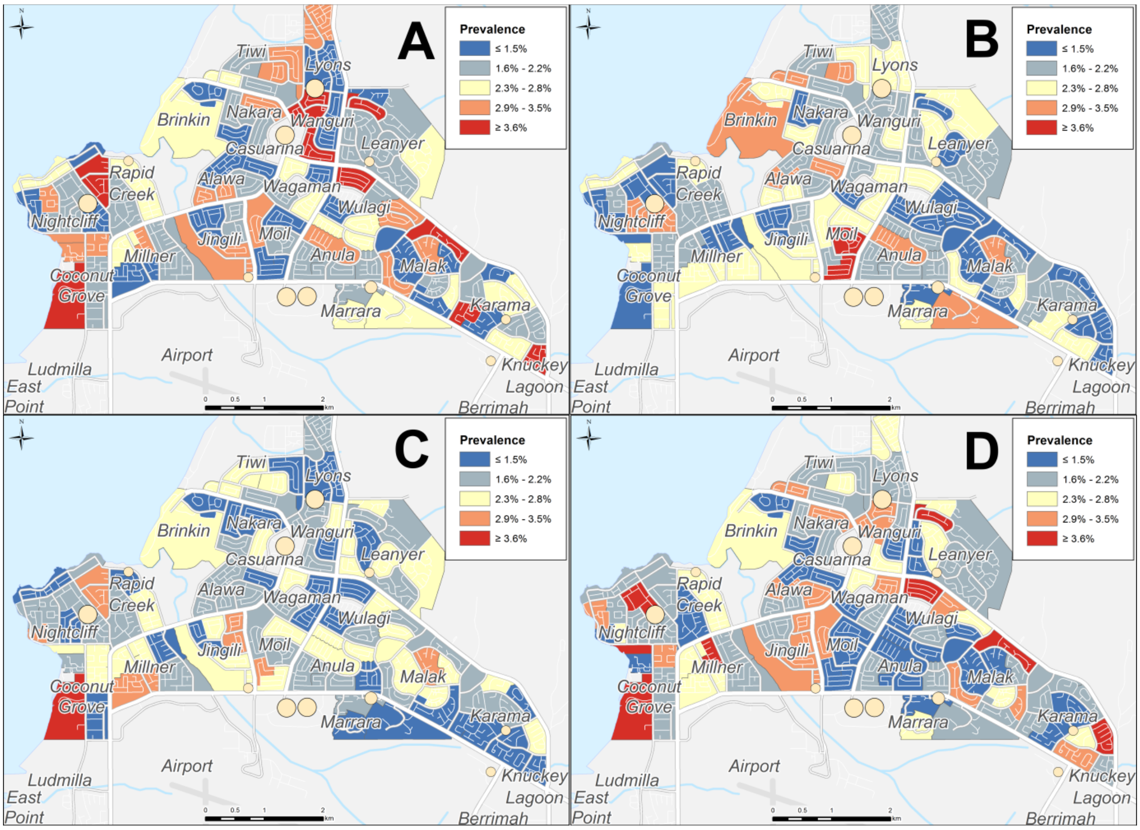

Figure 6.2: Maps of estimated prevalence of problem gambling in part of the study area generated from four spatial microsimulation models. Panel A shows predictions from the socio-demographic model. Panel B shows predictions from the environmental risk model. Panel C shows predictions from the gambling involvement model. Panel D shows predictions from the combined model.

6.6 Discussion

6.6.1 Interpretation of key results

This study has demonstrated how problem gambling prevalence estimates for small areas can be empirically-derived by combining survey data with census and health behavioural data. It has two key findings. First, extending the approach of Cullinan *et al.* (2012), it has shown how spatial microsimulation can incorporate constraints beyond socio-demographic variables and include measures of environmental risk factors and health behaviours. Logistic regression analysis suggests that the combination of health-behavioural measures and socio-demographic variables has the greatest explanatory power in terms of predicting health outcomes. In the case of problem gambling, the addition of a behavioural measure (gambling involvement) to a socio-demographic model increased the pseudo R^2 from 0.14 to 0.33. This demonstrates the potential afforded by the incorporation of health behavioural measures. Second, this study has shown that spatial microsimulation studies are highly sensitive to model specification.

Different model specifications can produce markedly different spatial patterns of the outcome variable of interest.

This study demonstrates the utility of marshalling administrative data on EGM expenditure in venues to predict gambling involvement rates in small geographic areas. Figure 6.1 demonstrated that a remarkably large proportion of the variance in mean per capita expenditure at the SA1 level, estimated using a Huff model, can be explained by survey-derived estimates gambling involvement for those same small area ($R^2 = 0.78$). Put plainly, this means that with the aid of a Huff model and venue-level EGM expenditure data, the number of people in a particular location gambling at a high intensity can be predicted with a remarkable degree of accuracy. Such a finding provides further evidence for the validity of Rose and Day's (1990) total consumption theory to the study of EGM expenditure (cf. Hansen & Rossow 2008; Lund 2008; Markham, Young, et al. 2014).

It is perhaps to be expected that health behavioural measures improve the predictive power of models of health outcomes. After all, in most cases health behaviours have greater causal proximity to outcomes than do environmental risk factors or socio-demographic variables. It is surprising, therefore, that health-behavioural variables have rarely been incorporated into spatial microsimulation studies as constraints. One reason for this absence is the usual omission of health-behavioural measures from census data. This study contributes to the literature by demonstrating how point-structured administrative data can be converted into constraints for small areas and then used for the purpose of spatial microsimulation.

The specification dependency among spatial microsimulation model results warrants further investigation. What is especially surprising is that the combined model, which incorporates the measure of gambling involvement, the single most explanatory variable, produced small-area estimates that were uncorrelated with those produced on the basis of gambling involvement alone. This might be explained by the optimisation goal of minimising total absolute error in a context of low multicollinearity. The combined model specification contained three socio-demographic variables but only one gambling involvement variable. Because total absolute error counts deviations from each constraint category equally regardless of the variable's explanatory power, the inclusion of a greater number of socio-demographic variables might 'weight' the analysis toward the constructs represented by the greatest number of constraint variables. A similar observation has been made in the case of cluster analysis (Hair et al. 2009), although the difference in the case of spatial microsimulation is that this problem is likely to be mitigated – not exacerbated – by the use of multicollinear predictor variables.

The feature selection weighting effect warrants future research into both its impact on results and into methods for mitigating it. One potential mitigation measure may be to duplicate a constraint that is under-weighted, thereby doubling its contribution to calculating total absolute error.[4] This approach might be generalised through the inclusion of arbitrary feature weights in total absolute error calculations in the combinatorial optimisation process. In this case, weights could be defined on the basis of principal components analysis or other methods of feature reduction. These suggested modifications to spatial microsimulation methods warrant future research, but are beyond the scope of this study. Until methods are developed for dealing with the feature selection weighting effect, analysts using spatial microsimulation should specify their models with a great deal of care and on the basis of theoretical concerns as well as goodness-of-fit indices.

The specification dependence exhibited in these results appear to be of more concern than those discussed previously in the peer-reviewed literature (e.g. Smith et al. 2009). Such variation deriving from sensitivity to model specification is not accounted for in recent methods developed to quantifying uncertainty in spatial microsimulation estimates (Nagle et al. 2014; Whitworth et al. 2016). While model specification uncertainty is by no means unique to spatial microsimulation, the results suggest that the problem may be especially acute when using this method. Model averaging provides one promising avenue by which model specification uncertainty may be quantified. The implication of sensitivity to model specification is that users of spatial microsimulation need to exercise great caution in ensuring that results are robust to variations in model configuration. This is especially important when, as in this case study, there is no 'gold standard' data against which external validation can take place. We believe that this finding is likely to be generalizable across geographic locations and problem domains in

[4] The authors are indebted to an anonymous reviewer for this suggestion.

cases when the outcome variable of interest is only moderately correlated with the predictor variables, as is very frequently the case (e.g. Anderson 2007; Whitworth et al. 2016). The improvement of model fit through the addition of health-behavioural measures is likely to be especially valuable in such situations.

6.6.2 Conclusions

The purpose of this paper was to explore the benefits of including health-behavioural variables in spatial microsimulation studies of health outcomes, with specific reference to problem gambling and gambling involvement. The study has made four contributions to the literature. First, it found that including health behavioural variables in a spatial microsimulation analysis was not only viable, but dramatically improved the explanatory power of related statistical models. This approach to incorporating auxiliary information should be encouraged in future applications of these methods. Second, the study demonstrated the accuracy of predicting gambling involvement in small areas on the basis of EGM expenditure data reported for individual gambling venues. Specifically, it found that the proportion of residents in small areas reporting high-gambling involvement in a survey could be accurately predicted on the basis of administrative data regarding gambling expenditure. Third, the inclusion of a health behavioural variable also demonstrated that spatial microsimulation results are dependent on model specification to an extent not generally appreciated in the literature. In cases where external validation against gold-standard data is not possible, sensitivity to model specification should be explicitly investigated. In cases with high sensitivity to model specification, results should be interpreted with caution. Future research might usefully develop and evaluate methods for assigning arbitrary weights to constraints when in the combinatory optimisation process. Finally, this study has provided four sets of empirically-calibrated estimates of problem gambling prevalence in small areas. It demonstrates that a great degree of spatial inequality exists in the prevalence of problem gambling, an inequality that is not only of concern in its own right, but also plays a role in furthering disparities among other economic and health outcomes. Such inequalities demand urgent policy attention.

Chapter 7: The relationship between electronic gaming machine accessibility and police-recorded domestic violence: A spatio-temporal analysis of 654 postcodes in Victoria, Australia, 2005-2014

7.1 Foreword

This chapter represents a deliberate break from the approach to understanding the health geography of gambling impacts adopted in earlier chapters. In the earlier chapters, 'problem gambling' figured prominently as the outcome variable of interest. This was partly a result of the availability of data measuring problem gambling using various psychometric screens, but also reflects the dominance of the construct of 'problem gambling' in the gambling studies literature (Cosgrave 2010; Reith 2007; Young 2013; Miller et al. 2016). However, the construct of problem gambling has many analytical flaws for public health research. Most importantly, screening instruments that measure problem gambling are effectively designed to measure subclinical gambling addiction (Svetieva & Walker 2008). While addiction undoubtedly increases gambling consumption, it is only one among many gambling-related harms, in the same way that alcohol dependence is only one among many alcohol-related harms. Consequently, gambling studies has tended to neglect the study of other gambling-related harms, which have only recently been rigorously enumerated and are poorly understood (Browne et al. 2016; Langham et al. 2016).

To that end, this chapter sets out a line of health geographic research that investigates the impact of EGM gambling without measuring 'problem gambling'. The study adopts methods from the study of the effects of alcohol-outlet density and applies them to gambling venues. Specifically, this chapter examines the spatio-temporal relationship between EGM accessibility and domestic violence, an under-researched gambling-related harm. It aims to provide an example of the possibilities that are open to geographical research into gambling when the construct of 'problem gambling' is no longer prioritised.

This chapter was published as:

Markham, F., Doran, B. & Young, M., 2016. The relationship between electronic gaming machine accessibility and police-recorded domestic violence: A spatio-temporal analysis of 654 postcodes in Victoria, Australia, 2005–2014. *Social Science & Medicine*, 162, p.106–114.

7.2 Abstract

An emerging body of research has documented an association between problem gambling and domestic violence in a range of study populations and locations. Yet little research has analysed this relationship at ecological scales. This study investigates the proposition that gambling accessibility and domestic violence rates in postcodes might be linked.

This study describes the association between police-recorded domestic violence and electronic gaming machine accessibility at the postcode level. Police recorded family incidents per 10,000 and domestic-violence related physical assault offenses per 10,000 were used as outcome variables. Electronic gaming machine accessibility was measured as electronic gaming machines per 10,000 and gambling venues per 100,000. Bayesian spatio-temporal mixed-effects models were used to estimate the associations between gambling accessibility and domestic violence, using annual postcode-level data in Victoria, Australia between 2005 and 2014, adjusting for a range of covariates.

Significant associations of policy-relevant magnitudes were found between all domestic violence and EGM accessibility variables. Postcodes with no electronic gaming machines were associated with 20% (95% credible interval [C.I.]: 15%, 24%) fewer family incidents per 10,000 and 30% (95% C.I.: 24%, 35%) fewer domestic-violence assaults per 10,000, when compared with postcodes with 75 electronic gaming machine per 10,000. The causal relations underlying these associations are unclear. Quasi-experimental research is required to determine if reducing gambling accessibility is likely to reduce the incidence of domestic violence.

7.3 Introduction

7.3.1 Background and rationale

A multitude of studies worldwide clearly demonstrate that domestic violence is an important public health issue (see Krantz 2002; Vine et al. 2010). Around 30% of women aged 15 or older worldwide are estimated to have been victims of intimate partner violence during their lifetime (Devries et al. 2013). In Australia, rates of domestic violence are above the global average, with 37% of women reporting ever experiencing domestic violence, and an estimated 12-month prevalence rate for intimate partner violence of 2.4% (Australian Bureau of Statistics 2012). A substantial portion of the global burden of ill health is attributable to domestic violence (Campbell 2002), with intimate partner violence alone estimated to account for 1% - 2% of disability adjusted life years for women globally (Lim et al. 2012).

Despite scholarly debate about the role of gender in domestic violence, it is clear from the overwhelming weight of evidence that domestic violence is mostly perpetrated by men against women (Hamby 2014). Cross-national evaluations of violence against women suggests that gender inequality is the most important predictor of the prevalence of such violence at the country level (Heise & Kotsadam 2015). However, an ecological understanding of domestic violence suggests that while male dominance must be at the foundation of any theoretical account, the aetiology of domestic violence is complex and multidimensional, with many factors influencing the probability of violence operating at several geographic scales and conceptual levels (Heise 1998).

An emerging body of research has examined the ecological factors that are related to the prevalence or incidence of domestic violence (Beyer et al. 2015; Pinchevsky & Wright 2012). Disadvantage has most frequently been found to be correlated with an elevated prevalence of domestic violence, whether conceptualised as ‘concentrated disadvantage’ in the framework of social disorganisation theory (e.g. Gracia et al. 2015; Pinchevsky & Wright 2012) or in terms of relative socioeconomic status (Beyer et al. 2015). Contrary to the expectations of social disorganisation theory, neither residential turnover nor neighbourhood ethnic heterogeneity have been consistently associated with the ecology of domestic violence (Pinchevsky & Wright 2012). For both of these variables, positive, negative and null correlations have been found. Other neighbourhood-level variables found to be associated with rates of domestic violence include collective efficacy, crime, disorder, gender inequality, rurality and alcohol-outlet density, although the research field is yet to yield a consensus on the importance of many of these contextual risk factors (Beyer et al. 2015).

Gambling is an under-researched contextual factor associated with domestic violence. Although an emerging body of research suggests a strong link between domestic violence and problem gambling at the individual level, domestic violence has too often remained a “hidden” issue for those researching the health and social impacts of gambling (Korn & Shaffer 1999). To date, the authors are aware of only one nationally-representative study that has examined the association between domestic violence and problem gambling. Re-analysing the US National Comorbidity Survey Replication (n = 3334; 18 years and older), Afifi and colleagues (2010) found that problem gambling was associated with both intimate partner violence perpetration and victimisation. In particular, endorsing five or more DSM-IV criteria for pathological gambling was associated with elevated odds of perpetrating severe marital violence (odds ratio [OR] = 20.4) and severe child abuse (OR = 13.2).

Most research investigating the association between domestic violence and gambling has investigated convenience samples recruited in clinical contexts or among those undertaking court-mandated programs. Dowling et al. (2016) performed a meta-analysis of 14 studies examining problem gambling among victims and perpetrators of intimate partner violence and found that among problem gamblers, 38.1% report victimisation and 36.5% report perpetrating violence. The prevalence of problem gambling was estimated to be 11.3% among perpetrators of intimate partner violence, compared with general population levels of 0.5% - 7.6% (Williams et al. 2012). In these studies, ‘problem gambling’ was measured using validated screening instruments like the Problem Gambling Severity Index or used the accepted diagnostic criteria directly (Dowling, Suomi, et al. 2016). The balance of evidence to date suggests an association between problem gambling and intimate partner violence that persists after adjusting for factors such as poor mental health.

It is plausible that these associations reflect a causal relationship between domestic violence and gambling, a proposition that is supported by the views of victims of violence whose partners have gambling problems (Muelleman et al. 2002; Suomi et al. 2013). For example, in one study of women aged 19 to 65 years presenting to an emergency department in Nebraska, 24% of IPV victims thought that their partner had a gambling problem. Among this subsample, 62% thought that their partners' gambling and violence were related (Muelleman et al. 2002). However victims' understanding of the causal direction of this relationship was not recorded.

Several studies suggest that gambling may increase the risk of violent incidents within families, as gambling may decrease perpetrators' feelings of control over their own and their families' lives (Korman et al. 2008; Suomi et al. 2013). For example, some of the interviewees in a study by Suomi et al. (2013) were assaulted when their partners returned home after a gambling session, angered about lost money. Other victims of violence in the same study reported that their own gambling losses triggered their partners' aggression. Such evidence does not imply that gambling is the ultimate cause of domestic violence. Rather, gambling here may serve as an indirect mediating factor that increases the frequency and severity of aggression in violent relationships.

In the reverse causal direction, gambling may also be used as a coping mechanism by victims and perpetrators to deal with domestic violence. For example, Blaszczynski and Nower (2002) suggest that for some problem gamblers, gambling serves as a method to "modulate affective states and/or meet specific psychological needs." In other words, for some people, gambling serves as a method for blocking out negative thoughts and memories (Wood & Griffiths 2007; Woolley & Livingstone 2010). This mode of gambling is highly prevalent, with around 19% of Australian adults reporting that they use gambling as a method of stress management (Australian Psychological Society 2015). We expect, therefore, a heightened gambling participation rate among both perpetrators and victims of violence as a result of violence-related stress.

Finally, it is likely that some portion of the correlation between domestic violence and problem gambling in previous studies results from shared risk factors that have not been controlled for. As such, the causal relationships between domestic violence and gambling are likely to be complex and multidirectional.

Given the documented associations between gambling and violence among individuals and their partners, it seems likely that neighbourhoods with highly accessible gambling opportunities will also have elevated rates of domestic violence. Gambling accessibility is of interest, as limiting venue accessibility is a policy lever that could potentially be used to reduce the incidence rate of domestic violence. Simple availability theory, borrowed from the alcohol research literature (Bruun et al. 1975; Stockwell & Gruenewald 2003), predicts that increased gambling accessibility will increase the rate of domestic violence, via the following causal chain: (i) as gambling accessibility increases, so too does gambling participation, frequency and intensity (Welte, Wieczorek, et al. 2004; Pearce et al. 2008; Young et al. 2012b), (ii) as total gambling increases, so too does the prevalence of individuals gambling at a harmful level (Rose & Day 1990; Grun & McKeigue 2000; Lund 2008; Markham, Young, et al. 2014; Markham et al. 2016), and (iii) as the number of people gambling more than they can afford in a population increases, so too does the incidence of domestic violence (Dowling, Suomi, et al. 2016). Alternatively, it is plausible that the causal chain may run in the other direction: (i) in areas with elevated rates of domestic violence, people seeking a cognitive escape will gamble more frequently and intensely than might otherwise be the case (Dowling, Suomi, et al. 2016) (ii) as gambling participation, frequency and intensity rise, gambling businesses become more profitable, and (iii) gambling operators will move or trade their licenses to areas with greater incidence of violence in order to profit from violence-induced demand (Marshall & Baker 2002).

Yet few studies have examined the link between gambling accessibility and domestic violence at the ecological level. Electronic gaming machine (EGM) accessibility is of particular interest, as EGMs are most closely associated with problem gambling (Welte, Barnes, et al. 2004; Productivity Commission 2010; MacLaren 2015) and are responsible for the largest amount of money spent by problem gamblers in Australia (Productivity Commission 2010). In Australia, three studies have touched on this relationship although none have specifically analysed domestic violence rates as distinct from assault rates more generally. Barratt and colleagues (2014) conducted a cross-sectional spatial analysis of gambling help and violence in local government areas of Victoria. Their study did not distinguish between domestic violence and other incidents of violence, but did find that violence and access to gambling help were significantly correlated. Similarly, Wheeler et al. (2010) conducted a statistical-local area level cross-sectional spatial analysis of crime in Victoria and found significant relationships between gambling losses and the population rate of non-income generating offenses, a category which includes domestic violence assault. Finally, in a separate study of statistical-local areas in South Australia, Wheeler and colleagues (2008) found no relationship between gambling losses and the rate of non-income generating offenses.

In the United States, a large and controversial body of econometric literature has examined the relationship between crime and casino accessibility at various ecological scales. These studies, comprehensively reviewed by Walker (2010), have examined crime at county and city scales and have had mixed results, with some finding associations between various offenses and casino availability and others finding no significant relationship. Few of these ecologic studies specifically examined domestic violence with any degree of rigour. Other studies in the United States have been unable to disaggregate domestic violence from other violence due to their reliance on the classification of offenses imposed by the Uniform Crime Reports' standard format (e.g. Evans & Topoleski 2002; Grinols & Mustard 2006; Reece 2010).

In summary, the extant literature has examined the relationship between problem gambling and domestic violence at the individual level only and in general has found significant relationships. Ecological studies of the impact of gambling accessibility on crime have had mixed results and have rarely focused on domestic violence specifically. This relative scholarly neglect is perhaps unsurprising, as both domestic violence (DeVerteuil 2015) and gambling (Marshall & Baker 2001a) have been subject to relatively little geographic research.

7.3.2 Objectives

In this study, we examine the relationship between domestic violence and EGM accessibility at the neighbourhood level. Specifically, we undertake a spatio-temporal analysis of EGM accessibility in Victoria and annual police-reported domestic violence between 2005 and 2014 at the postcode level. We estimate the association between the incidence of domestic violence and EGM accessibility, using two measures of EGM accessibility and two measures of police-recorded domestic violence. Because both EGM accessibility and domestic violence are correlated with variables such as socio-economic disadvantage that vary spatially and temporally, we adjust for a range of covariates that differ between postcodes.

7.4 Methods

A Bayesian spatio-temporal Poisson regression approach was used to model the relationship between domestic violence, EGM accessibility and socio-demographic covariates in Victoria, using police-recorded data aggregated by postcode and year from 2005-2015. Bayesian spatial or spatio-temporal models are appropriate for investigating ecological correlations between variables under conditions of space-time dependence. While Bayesian methods are widely used in epidemiology, they have rarely been applied to research on domestic violence (Cunradi et al. 2011; Freisthler & Weiss 2008; Gracia et al. 2014).

7.4.1 Outcome variables

We used two measures of domestic violence, both derived from the Law Enforcement Assistance Program (LEAP) database used by the Victoria Police to store police records, and provided by the Crime Statistics Agency of Victoria. The first measure was the number of 'family incidents' recorded by police in a postcode in a given calendar year. A family incident is recorded whenever domestic-violence assault, interfamilial sexual offences or child abuse are reported to police, including incidents that do not lead any charges being laid (Victoria Police 2014). The number of family incidents recorded grew dramatically over the study period, from 28,424 in 2005 to 68,091 in 2014. The increase in recorded family incidents accelerated after 2011 when new police reporting guidelines were introduced, moving from an average annual growth rate of 5.7% from 2005-2010 to an average of 16.4% from 2011-2014.

Due to the dynamism of the family incidents measure, the number of domestic violence assault offenses was used as a secondary outcome variable. Domestic violence assaults were recorded when charges were laid by police for 'assault and related offences' (classification A20) with respect to a family incident as defined above. While this category only covers some types of domestic violence acts and these for only a subset of cases, the category was selected as it was the most numerous family-incident related crime against the person. Additionally, the number of recorded family-incident related physical assaults grew less rapidly over the study period than other family-incident related offenses. Recording of domestic-violence assaults increased by 286% between 2005 and 2014, compared with a 603% growth rate in recording of all other family-incident related offenses. Counts of police-recorded domestic violence assaults were also aggregated to the postcode level.

Both measures of domestic violence include incidents where the perpetrators are both men and women. In 2014-15, the person identified by police as the ‘primary aggressor’ (Victoria Police 2014) was a male aged 15 or more in 76% of family incidents. Women and children aged 14 or younger were the victims (or ‘affected family members’) in 78% of incidents.

7.4.2 Predictor variables

Data on the residential population of each postcode was required to calculate rates of violence. It is worth noting that while much of the American literature reviewed by Walker (2010) is concerned with adjusting for the effect of non-residents visiting in order to gamble at casinos, this issue is unlikely to have relevance in the Victorian context due to the dispersed nature of EGMs across the community. Annual estimated residential populations (ERPs) for Statistical Areas 2 (median 2011 population of 8,562) were obtained from the Australian Bureau of Statistics for the study period (Australian Bureau of Statistics 2015) and were converted to postal areas. The conversion took place using a custom-made concordance file. The concordance was produced by weighting calculating the proportion of each SA2’s census count was resident in each postal area, using the SA1 unit to produce such an estimate. SA2 population ERPs were then multiplied by these postal-area specific proportions, yielding an estimated allocation of SA2 populations to postal areas. These allocations were then summed for all postal areas.

The predictor variable of interest was EGM accessibility. Accessibility was measured the postcode level in two ways: EGMs per 10,000 persons and EGM venues per 100,000 persons, each recorded separately for each postcode during each calendar year. While venue density may be considered a better measure of accessibility as it may be that distance to nearest venue is more important than venue size, EGM density is also of interest as it combines the venue density metric with an indicator of capacity. Data detailing the number of EGMs in each venue in Victoria during each year of the study period were obtained from the Victorian Commission for Gambling and Liquor Regulation. The postcode of each venue was extracted from its street address. EGM density was calculated by summing the number of EGMs in each postcode for each calendar year, and dividing that by the postcode estimated residential population.

The remaining predictor variables were derived from the Australian Census of Population and Housing, which took place in 2001, 2006 and 2011. Linear interpolation was used to provide estimates of census variable in intercensal years, while 2011 values were carried forward for the years 2012–2014.

Following calls to measure the health effects of gender (e.g. Phillips 2011), we measured economic gender inequality by calculating the percentage of total personal income in the postcode that accrued to women. We hypothesised that domestic violence rates would be lower in areas where women had relatively more economic autonomy, as measured by the female share of personal income. We estimated the female income share using self-reported personal income data from the census, coding income brackets according to their midpoints and treating the top-coded bracket following Fleming and Measham's (2015) assumption that income is Pareto distributed.

As economic disadvantage is another key variable that is associated with both recorded domestic violence (Beyer et al. 2015; Pinchevsky & Wright 2012) and EGM density (Marshall & Baker 2001a; Marshall & Baker 2002), the Socio-Economic Index for Areas (SEIFA) Index of Economic Resources (IER) was included as a postcode-level measure of poverty and prosperity. The IER is a composite measure of various indicators of wealth and income, which means that a more complete picture of economic status can be gained by using a single indicator, without introducing multicollinearity among predictor variables (Australian Bureau of Statistics 2008).

Other included predictor variables were: the percentage of residents in the postcode who speak only English, a variable relevant to social disorganisation theory; the child-to-woman ratio, selected because the number of children present in the home is an important risk factor for victimisation (Stith et al. 2004), measured as the number of children aged 0–4 divided by the number of women aged 15–45; the percentage of the postcode who identify as Indigenous, selected due to the higher rates of victimisation among Indigenous women in Australia (Mouzos & Makkai 2004; Al-Yaman et al. 2006); the median age of residents in the postcode, selected because domestic violence risk decreases with age (Stith et al. 2004); and the Open Accessibility and Remoteness Index for Australia (ARIA+), a continuous measure of the geographic accessibility of locations in Australia, selected to capture any urban-rural differences in police recording rates or domestic violence rates (Markham 2015). The index ranges from 0 in the most urban areas to a maximum of 15 in the most isolated parts of Australia.

During exploratory modelling, each predictor variable was tested in separate Poisson regression models, using no transformation, logarithmic transformation and a second-order polynomial transformation. The variable form that yielded the lowest DIC was retained for the main analysis.

7.4.3 Spatial units

Due to changes in the configuration of postcodes during the study period, four postcodes were removed from the analysis. In addition, nine postcodes were removed as they represented non-residential areas including university campuses (n = 3), military bases (n = 3), industrial zones (n = 2) or the central business district (n = 1). These areas were excluded for reasons of missing data or to reduce potential bias induced by the high ratio of visitors to residents in the central business district. This left 654 postcodes with 10 years of annual data remaining in the analysis. Postcode boundaries were approximated using the Australian Bureau of Statistics' Postal Areas.

7.4.4 Statistical methods

Exploratory Poisson regression modelling was undertaken assuming that data were identically and independently distributed. Variance-inflation factors (VIFs) were checked to ensure that multicollinearity was not problematic. Because the logarithmic transformations of venue density and EGM density are highly correlated ($r = 0.96$), they could not be included simultaneously in the same model, leading to the creation of separate models. All other VIFs were in the range 1.2–3.5 (pairwise correlation coefficients between variables are listed in Appendix H, Table H.1). As expected, Moran's test revealed that significant spatial autocorrelation remained in the models' residuals, while unambiguous temporal trends were also evident in model residuals.

Consequently, a spatio-temporal analysis was required in order to incorporate spatial and temporal smoothing into the model.

A Bayesian Poisson mixed effects model was therefore used to estimate the associations between domestic violence, EGM density and socio-demographic correlates, taking into account the spatio-temporal autocorrelation non-parametrically. A modelling approach drawing on that of Knorr-Held (2000) was adopted, as it allows for non-separable spatial and temporal autocorrelation structures. As we were interested examining fixed estimates of parameter values rather investigating than the random effects themselves, Ugarte et al.'s (2012) simplification that eliminates several nuisance parameters from the Knorr-Held specification was adopted. The model used in the spatio-temporal analysis was:

$$Incidents_{kt} \sim \text{Poisson}(\mu_{kt},\ pop_{kt})$$

$$\log(\mu_{kt}) = \beta X_{kt} + \log(pop_{kt}) + \phi_t$$

$$\beta \sim \text{N}(0, 0.001)$$

$$\phi_1 \sim \text{N}(0, \tau^2 Q(W,\rho)^{-1}) \qquad t = 1$$

$$\phi_t | \phi_{t-1} \sim \text{N}(\gamma\phi_{t-1}, \tau^2 Q(W,\rho)^{-1}) \qquad t = 2, \dots, T$$

$$Q(W,\rho) = \rho[\text{diag}(W1) - W] + (1-\rho)I$$

$$\tau^2 \sim \text{Inverse Gamma}(0.001, 0.001)$$

$$\rho, \gamma \sim \text{Uniform}(0, 1)$$

where the study site is divided into $k = 1, \dots, K$ postcodes and $t = 1, \dots, T$ years; $Incidents_{kt}$ records the number of recorded family incidents in postcode k in year t; pop_{kt} records the population of postcode k in year t, included as an offset term to account for variations in postcode size; X_{kt} is a matrix of predictor variables in postcode k in year t and β is a commensurate vector of estimated regression coefficients; ϕ_t captures spatio-temporal correlation through a set of random effects. Temporal autocorrelation is modelled through a first order autoregressive term, while spatial autocorrelation enters the model through the random effects precision matrix. The precision matrix is given by $Q(W,\rho)$ where 1 is a vector of ones and I is the $K \times K$ identity matrix. The spatial adjacency matrix is given by W. ρ and γ vary, respectively, according to the degree of spatial and temporal autocorrelation in the data, with 0.0 indicating no autocorrelation and 1.0 indicating strong autocorrelation.

The adjacency matrix was calculated using queen's contiguity in R using the *spdep* package (Bivand et al. 2008) with a tolerance of 10 meters. The model was fitted using the *CARBayesST* package (Lee et al. 2015). Model parameters were estimated using Markov-Chain Monte Carlo (MCMC) simulations. A burn in period of 40,000 iterations was used, after which 200,000 samples were made. A thinning factor of five was used to reduce correlation among the simulations, meaning that parameters were effectively estimated from 40,000 simulations. Convergence was assessed by examining trace plots, with estimates of all the parameters of interest converging well.

7.5 Results

7.5.1 Descriptive statistics

After excluding ineligible postcodes, 4,970,000 persons were resident in the study area in 2005, rising to 5,800,000 in 2014. The exclusion of ineligible postcodes removed less than 1% of the Victorian population from the analysis.

Among Victorian postcodes in 2005-2014, the median rate of family incidents and domestic-violence assaults per 10,000 were 52 and 10, respectively (see Table 7.1). However, the distribution of incidents was highly positively skewed, with the mean rate of family incidents and domestic-violence assaults much higher at 65 and 17, respectively. The distribution of EGM accessibility variables were similarly positively skewed.

Table 7.1: Descriptive statistics summarising recorded domestic violence, electronic gaming machine density and socio-demographic covariates in 654 Victorian postcodes, each year from 2005-2014

	Mean	Median	SD	Min	Max
Total population	8220.4	2831.3	11422.6	123.3	93875.6
Family incidents per year per 10,000	65.1	52.2	57.5	0.0	754.0
Domestic-violence assaults per year per 10,000	16.5	10.3	23.1	0.0	475.1
Venues per 100,000	6.0	0.0	11.2	0.0	91.7
EGMs per 10,000	15.5	0.0	28.7	0.0	232.5
Index of economic resources	1004.1	999.1	65.0	811.1	1213.0
Female income share	37.8	37.8	4.7	13.7	77.8
Per cent only speak English	84.9	91.6	15.4	16.6	100.0
Child-to-woman ratio	0.3	0.3	0.1	0.0	1.6
Per cent Indigenous	0.9	0.5	1.2	0.0	18.1
Median age	41.1	40.6	5.8	20.0	65.0
Accessibility and remoteness index for Australia	0.4	0.3	0.4	0.0	2.1

Notes: Summary statistics presented in this table were not weighted by postcode population and therefore should be interpreted as representing averages among postcodes rather than the population average at state level.

Police-recorded domestic violence was highly concentrated in particular regions of Victoria (see Figure 7.1). In particular, high rates of violence were recorded in the northern, western and far south-eastern suburbs of Melbourne. Similarly, EGMs are concentrated in particular regions, particularly Melbourne’s western, north-western and south-eastern fringes.

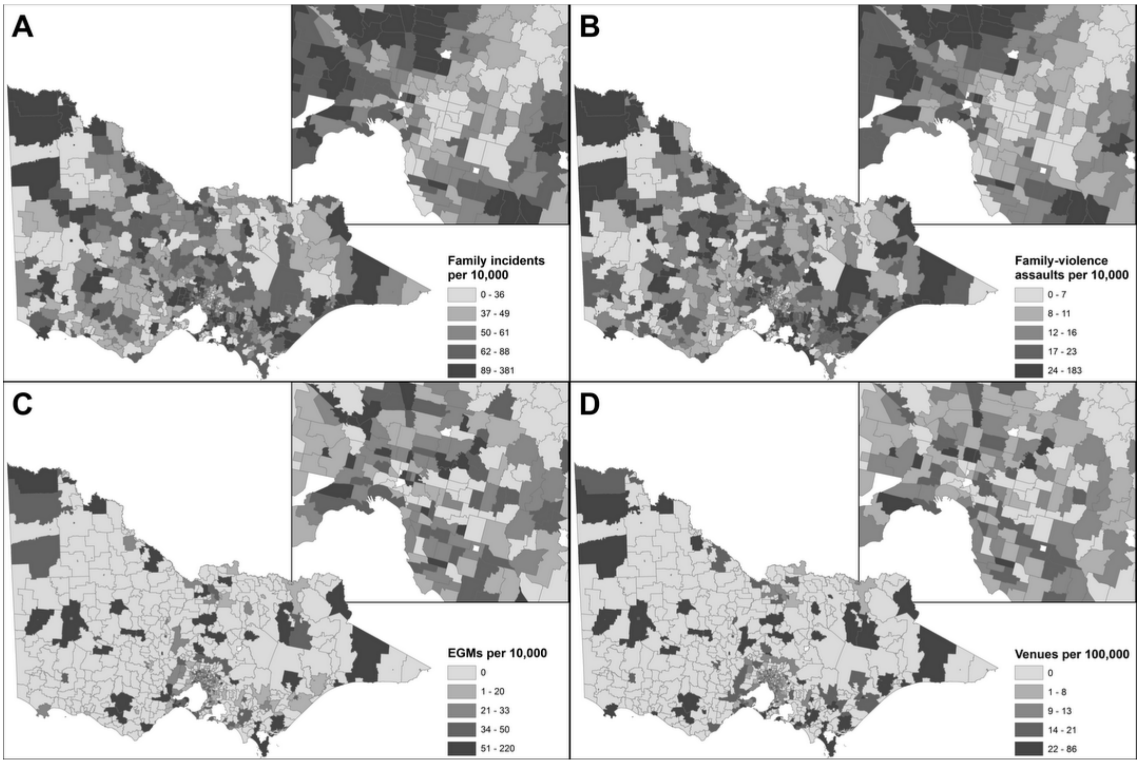

Figure 7.1: Map of the unsmoothed spatial distribution of postcodes in Victoria showing (A) family incidents per 10,000; (B) family-violence assaults per 10,000; (C) EGMs per 10,000, and (D) EGM venues per 100,000. Inset maps show Greater Melbourne. All maps show the mean of annual postcode rates.

7.5.2 Main results

Both venue density and EGM density were significantly associated with family incidents and domestic-violence assaults (Table 7.2). The model fitting process revealed the relationship to be best modelled as a logarithmic curve, indicating that the association increases rapidly with the first EGM venue in a postcode, followed by marginally smaller increases in domestic violence as accessibility increases thereafter. This correlation held both when controlling for other variables in multiple regression (Table 7.2) and in bivariate models (Appendix H, Table H.2).

Table 7.2: Associations between domestic violence, EGM accessibility and socio-demographic characteristics from multivariate Bayesian spatio-temporal analysis

	β coefficients							
	Family incidents				Domestic-violence assaults			
	Model 1		Model 2		Model 3		Model 4	
	Est.	95% C.I.	Est.	95% C.I.	Est.	95% C.I.	Est.	95% C.I.
Intercept	**-8.3**	**[-9.3, -7.3]**	**-8.2**	**[-9.2, -7.2]**	**-9.6**	**[-11.0, -7.9]**	**-9.6**	**[-11.1, -8.0]**
ln (Venues per 100,000 + 1) × 10^1	**0.8**	**[0.6, 0.9]**			**0.9**	**[0.6, 1.1]**		
ln (E.G.M.s per 10,000 + 1) × 10^1			**0.5**	**[0.4, 0.6]**			**0.8**	**[0.6, 1.1]**
I.E.R. × 10^3	**-5.5**	**[-6.0, -5.0]**	**-5.6**	**[-6.1, -5.1]**	**-5.8**	**[-6.6, -5.0]**	**-5.9**	**[-6.6, -5.3]**
I.E.R.^{2} × 10^1	-0.5	[-4.0, 2.9]	-0.5	[-4.2, 3.4]	4.1	[-1.3, 9.2]	4.2	[-1.1, 9.3]
ln Fem. income share	**0.6**	**[0.3, 0.8]**	**0.5**	**[0.3, 0.8]**	**0.6**	**[0.3, 0.9]**	**0.6**	**[0.3, 0.9]**
ln % English only × 10^1	**2.0**	**[0.5, 3.2]**	**1.9**	**[0.6, 3.2]**	0.9	[-1.0, 2.7]	0.9	[-0.9, 2.6]
Child-to-woman ratio × 10^1	**4.7**	**[1.3, 7.7]**	**4.4**	**[1.6, 7.6]**	**4.9**	**[0.5, 9.4]**	4.7	[-0.3, 9.2]
ln (% Indigenous + 1) × 10^1	**2.2**	**[1.6, 2.8]**	**2.2**	**[1.6, 2.9]**	**1.9**	**[0.9, 2.9]**	**1.8**	**[0.9, 2.7]**
Median age × 10^2	-0.4	[-0.9, 0.4]	-0.4	[-0.9, 0.2]	-0.5	[-1.5, 0.3]	-0.5	[-1.4, 0.3]
Median age^2	2.4	[-1.8, 7.5]	1.4	[-3.6, 6.7]	4.5	[-2.0, 12.0]	4.7	[-2.3, 11.7]
ln (OARIA + 1) × 10^1	**-6.6**	**[-10.3, -3.1]**	**-7.6**	**[-11.0, -4.9]**	-3.4	[-7.9, 0.8]	-3.5	[-7.0, 0.4]
τ^2	.11	[.11, .12]	.12	[.11, .13]	.23	[.20, .25]	.23	[.20, .25]
ρ	.95	[.92, .97]	.95	[.92, .97]	.97	[.94, .98]	.97	[.94, .98]
γ	.98	[.96, 1.00]	.98	[.97, 1.00]	.89	[.87, .92]	.89	[.87, .92]
D.I.C.	39352		39347		28705		28705	

Notes: Est. = estimate; C.I. = credible interval; E.G.M. = electronic gaming machine; I.E.R. = Index of economic resources; Fem. = female; OARIA = Open Accessibility and Remoteness Index of Australia; D.I.C. = Deviance information criterion. *ln* () indicates the natural logarithm. Bold type indicates coefficients for which the 95% C.I. does not contain zero. Model coefficients are estimated simultaneously from a multivariate spatio-temporal model containing all the parameters for which coefficients were listed.

The association between EGM density and domestic violence are presented in terms of absolute risk in Figure 7.2. It is clear that while EGM accessibility is associated with violence in all four models, it explains only a small fraction of the inter-postcode variation in domestic violence rates. Table 7.2 shows that alternating between venues per 100,000 and EGMs per 10,000 has little substantive effect on the estimated association between EGM accessibility and domestic violence, although the DICs reported in Table 7.2 suggest that use of venues per 100,000 marginally improves model fit.

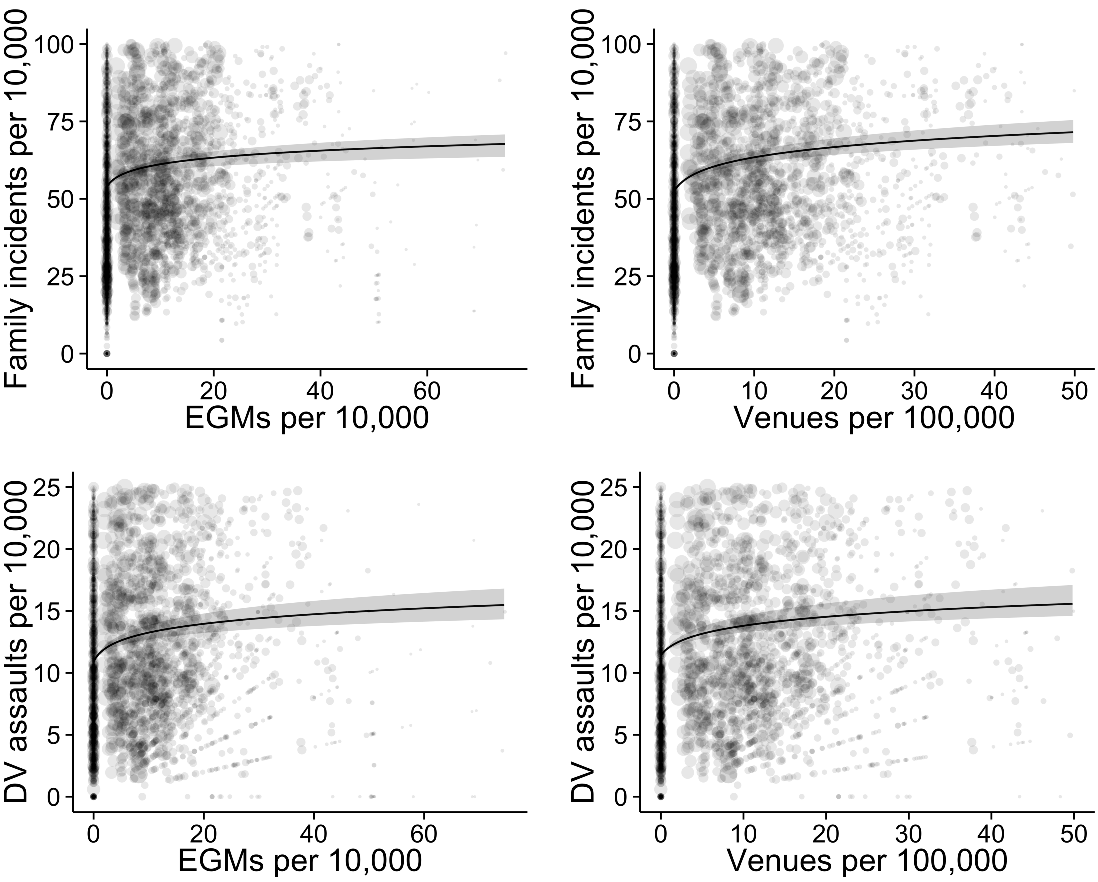

Figure 7.2: Associations between domestic violence outcomes and EGM accessibility in Victorian postcodes, 2005–2014, after adjusting for socio-demographic characteristics from Bayesian spatial analysis. The trendline shows the estimated association between EGM accessibility and domestic violence, with the shaded grey area representing the 95% credible interval. Each circle represents a single postcode in a single year, with the area of circles sized proportionate to the population of the postcode.

The index of economic resources was negatively correlated with both measures of domestic violence, meaning that there were more incidents of police-recorded violence per capita in poorer areas. Contrary to our expectations, the female income share was positively correlated with the rate of domestic violence, meaning that as women in a postcode gain relatively more economic power, recorded violence rates increase rather than decrease. The percentage of people only speaking English was negatively correlated with family incidents but was not correlated with domestic-violence assaults. One potential explanation of this may be that communication barriers prevent those from a non-English speaking background from reporting violence to police. The accessibility and remoteness index was negatively correlated with violence rates, meaning that police recorded a greater rate of family incidents in more urban areas.

Both ρ and γ were estimated to be relatively close to 1.0. This indicates a great degree of both spatial and temporal autocorrelation among the data. Because recording of family incidents increased during the study period much more rapidly than domestic-violence assaults, γ is correspondingly higher for family incidents.

7.6 Discussion and conclusions

7.6.1 Key results

EGM density was significantly associated with both the rate of family incidents and the rate of domestic-violence assaults (Table 7.2). As Figure 7.2 shows, EGM-free postcodes were associated with a mean incidence rate of 54 family incidents per 10,000 and 11 domestic-violence assaults per 10,000. The mean incidence rate for postcodes with 75 EGMs per 10,000 (approximately 2 SDs above the mean), was 68 family incidents per 10,000 (95% C.I.:64, 71) and 16 domestic-violence assaults per 10,000 (95% C.I.: 14, 17). Similar patterns were evident when venue density was used as the predictor variable. In terms of relative risk, postcodes with no EGMs were associated with 20% (95% credible interval [C.I.]: 15%, 24%) fewer family incidents per 10,000 and 30% (95% C.I.: 24%, 35%) fewer domestic-violence assaults per 10,000, when compared with postcodes with 75 EGMs per 10,000.

Other postcode-level covariates such as economic disadvantage, female income share, the percentage of households only speaking English, the number of children per woman, the proportion of Indigenous residents and urbanicity were also found to be correlated with the increased incidence of police-recorded domestic violence.

7.6.2 Limitations and generalisability

These findings are subject to several limitations. First, police recorded data routinely under-record domestic violence (Felson et al. 2002), with recorded cases sometimes likened to the tip of an "iceberg" of statistically-invisible, unrecorded cases (Gracia 2004). This is likely to downward bias the estimated incidence of domestic violence and thus the magnitude of the estimates of absolute risk. However, our estimates of relative risk will only be biased if recording rates are spatially correlated with EGM density. We cautiously suggest that there is little *a priori* reason to expect that this would be the case. Second, our research is subject to the modifiable areal unit problem (MAUP: Openshaw 1984). Postcodes are not the ideal spatial units for conducting this kind of analysis. Future research could use geocoded police data to analyse the relationship between EGM accessibility and domestic violence at fine spatial scales using frame-invariant statistics (Tobler 1990), thereby mitigating the MAUP. Analysis of geocoded police data would also enable the use of more sophisticated sub-zonal measures of EGM accessibility. Third, our estimates are associations only, limiting any causal inferences that can be made about the relationship between EGM density and domestic violence. Fourth, we did not set out to specifically investigate interactions between EGM density and other predictors of domestic violence but have instead estimated the average association between EGM density and violence across social space. Future research might investigate the potential socio-spatial non-stationarity of these relationships.

We see little *prima facie* reason to doubt that our results are generalizable to other jurisdictions in high income countries where EGMs are accessible throughout the community. While the incidence rates are highly dependent on police recording protocols, we expect that the underlying relationship will be present in other jurisdictions.

7.6.3 Interpretation

There is an association between EGM density and domestic violence rates among postcodes, above and beyond that explained by the geography of contextual factors like disadvantage. That is to say, once the effect of disadvantage and other covariates are controlled for, a correlation between EGM density and violence remains. The logarithmic shape of the relationship suggests that the first few EGM venues added to an area have the greatest impact on domestic violence rates.

The association between domestic violence and social disadvantage has been controlled for through the use of covariates. However, this study has not investigated the interaction between exposure to poker machines, disadvantage and domestic violence. It may be the case that the impact of poker machines is greater in poorer areas. Future analyses may wish to stratify their analyses by social disadvantage, or include it as an interaction term in the generalised linear model.

Gambling is likely to be both a cause and effect of domestic violence. Furthermore, despite our efforts to find neighbourhood-level proxies for domestic violence risk factors, it is likely that the correlation between EGM accessibility and domestic violence results partly from the correlation of both these variables with other unobserved variables. We expect that all three of these causal pathways contribute to some extent to the association we have documented, although their relative importance is unclear.

The magnitude of these estimated associations is larger than we anticipated and suggests that the relationship between EGM accessibility and domestic violence is policy relevant. If all of this association results from gambling-caused domestic violence (rather than the other two causal pathways mentioned above), we would expect that reducing the accessibility of EGMs or removing them entirely would result in a substantially reduced incidence of domestic violence. A more plausible interpretation is that only a fraction of the estimated association is attributable to gambling-caused domestic violence. Further study using strong quasi-experimental designs is needed to disentangle the multiple causal pathways to this association and to estimate the magnitude of potential reductions in the domestic violence incidence rate that may occur if EGM accessibility were decreased.

As Figure 7.2 demonstrates, EGM accessibility only explains a small portion of the variation in the rate of police-reported domestic violence between postcodes. This implies that reducing EGM density is unlikely to be a policy priority for those seeking to reduce rates of domestic violence. However, domestic violence should be considered as a relevant social impact by authorities licensing EGMs.

Consistent with the broader literature and social disorganisation theory, economically disadvantaged postcodes have an elevated rate of police-recorded domestic violence (Beyer et al. 2015; Pinchevsky & Wright 2012), although it is possible that this correlation is due to more intensive policing of poorer areas. Our finding that postcodes with greater linguistic diversity experience lower rates of police-recorded domestic violence is contrary to the expectations of social disorganisation theory, but consistent with non-ecological surveys in Australia (Mouzos & Makkai 2004; O'Donnell et al. 2002). Contrary to our expectations, postcodes with a greater female income share also had elevated rates of violence. Jewkes (2002), writing about similar relationships between intimate partner violence and female education in the United States and South Africa, suggests that as women gain more power they pose a greater challenge to patriarchal authority, a challenge that may be met with violence. Finally, the finding that police-recorded violence rates are higher in urban areas when urbanicity is measured using a continuous measure of remoteness and accessibility is novel in the research literature, which has rarely measured the urban-rural continuum using non-binary measures in developed countries (Beyer et al. 2015).

7.7 Conclusions

EGM accessibility is associated with police-recorded domestic violence incidence in postcodes in Victoria. Reducing EGM accessibility may potentially provide an avenue for reducing the incidence of domestic violence. Further research utilising strong quasi-experimental designs should be undertaken to disentangle the causal relations underlying this association.

Chapter 8: Conclusions

8.1 Summary of main research findings

This thesis consists of six research papers on the spatial distribution of the impacts of electronic gaming machines (EGMs) and the relationship between EGM losses and problem gambling. Detailed findings for each of these studies are described in their respective chapters. Here, I wish to briefly discuss the combined findings of this collection of studies in more general terms and outline the broader contribution of the research.

The first key finding is the conclusion that gambling losses, especially EGM losses, are a useful proxy for gambling-related harm at the population level. Positive correlations between these two variables were found for individuals, EGM venues and Australian states and territories. In the case of individuals, the shape of the dose-response relationship between EGM gambling losses and problem gambling risk was found to be either linear or r-shaped, rather than J-shaped as previously supposed. Venue-level and jurisdictional-level associations were consistent with a linear relationship (although curvilinear relationships were not specifically investigated). Taken together, these findings support the hypothesis posited by Total Consumption Theory (Rose & Day 1990) that increased levels of EGM gambling expenditure lead to increased levels of gambling-related harm. Furthermore, the evidence for a linear loss-risk relationship supports the appropriateness of what Chokshi *et al.* (2015, p.1339) call traditional public health measures 'around "reduce, restrict, limit, ban"' for EGMs.

The second set of research findings relates to the development of new methods for estimating the prevalence of problem gambling in small areas and for modelling the spatial behaviour of EGM gamblers. The study showed that EGM venue visitors are remarkably predictable with regard to their venue choice behaviour (R^2 = 0.64), with those who gambled on EGMs having even more predictable spatial behaviour (R^2 = 0.72). The ability to estimate gamblers' spatial behaviour enabled the production of empirically-calibrated small area estimates of problem gambling prevalence that take EGM losses at venues into account. Small area estimates of problem gambling prevalence that are based on EGM losses at venues produce quite different small area estimates to those that are constrained by socio-demographic variables only. To the best of the authors' knowledge, these are the first set of empirically-calibrated small area estimates of problem gambling prevalence that have been published. They constitute a substantial improvement on purely normative or predictive spatial models of problem gambling risk that have been published to date.

Third, the final chapter in the thesis took a different approach to the spatial estimation of gambling-related harm. Rather than seeking to estimate the spatial distribution of problem gamblers using survey data, it instead sought to estimate the spatial distribution of a specific gambling-related harm using a complete administrative dataset. In this specific case, the study found that a spatial association exists between the accessibility of EGMs and the incidence of police-recorded domestic violence. This association was robust regardless of whether the accessibility of EGMs was measured in terms of EGM density or EGM venue density, nor whether domestic violence was specified as any police-recorded family-violence related incident, or whether only domestic violence incidents that lead to a charge of assault being laid. The magnitude of the spatial association was surprisingly large, with postcodes with no electronic gaming machines having 20% (95% credible interval [Cr.I.]: 15%, 24%) fewer family incidents per 10,000 and 30% (95% Cr.I.: 24%, 35%) fewer domestic-violence assaults per 10,000, when compared with postcodes with 75 electronic gaming machine per 10,000. Gambling is likely to be both a cause and effect of domestic violence, with some of the estimated association also likely to derive from the correlation of both these variables with other unobserved variables.

8.2 Implications of findings

These findings have several implications for policy and research. First, the methods developed in this thesis can assist researchers and decision-makers to understand the spatial distribution of the negative impact of EGMs. In particular, decision-makers involved in social impact assessments may now have recourse to independent and reliable data regarding two crucial issues: (a) the catchments of EGM venues, and (b) the number of problem gamblers living in small geographic areas. In the case of EGM catchments, previous work (Doran et al. 2007) clearly demonstrated that simplistic rules-of-thumb (e.g. that the social impacts of an EGM fall within a 2.5 km radius; KPMG Consulting 2000) are inappropriate. While it was known that EGM venue catchments are variable between gambling venues, this thesis has provided calibrated tools for estimating their extent from desktop data. Clearly, in a social impact assessment process, it is imperative to understand the spatial extent of any social impacts, and this thesis has provided a method for estimating that extent. In the case of estimating the number of problem gamblers, this thesis has produced a method for estimating their numbers and residential location by combining survey, administrative and census data. This can assist social impact assessment by providing a baseline estimate of the existing number of problem gamblers in a local area. More importantly, if projections of the amount of money that will be lost on new EGMs are available, then the impact of additional EGMs can be estimated in terms of changes in the prevalence of problem gambling in specific geographic areas.

For researchers, the ability to estimate the location of EGM gamblers offers the possibility to understand the spatial distribution of gambling-related harm. This has the potential to assist in answering several research questions regarding the social distribution of the burden of gambling-related harm. For example, studies of the health impacts of air pollution have found that disadvantaged groups suffer from a 'triple jeopardy' of (a) a greater exposure to air pollution, (b) a greater burden of reduced health from social factors, and (c) a detrimental interaction between the social determinants of health and exposure to air pollution, resulting in a concentration of serious health impacts (Jerrett et al. 2009). The methods developed in this thesis offer the possibility to evaluate whether or not such a triple jeopardy exists for EGM gambling. Similarly, these methods offer the ability to contribute to the understanding of the relationship between community-level and individual-level risk for EGM gambling. Much previous research has found that both EGMs (e.g. Marshall & Baker 2002) and EGM expenditure (e.g. Rintoul et al. 2013) are concentrated in the poorest areas of cities like Sydney and Melbourne. However, as Delfabbro and King (2017) point out, problem gambling prevalence studies in these same cities generally do not show a significant correlation (or show an inverse U-shaped correlation) between personal income and individual-level problem gambling risk. These observations have been used to support the proposition that the spatial concentration of EGM venues in poorer areas does not translate into a corresponding concentration of gambling-related harms among disadvantaged people (Delfabbro & King 2017). The new methods presented in this thesis offer the possibility to unpack the processes that produce this apparent contradiction.

Second, the final study in this thesis which evaluated the relationship between EGM accessibility and police-recorded domestic violence has several key implications. In terms of policy, this study suggests that domestic violence impacts should be considered when regulators make decisions about granting licenses for poker machines. While we were unable to draw causal conclusions, a preventive approach suggests that licensing decisions should be informed by the understanding that increased EGM accessibility may lead to increased domestic violence incidence. Indeed, this association (and others like it) have wide ranging implications for understanding the social costs of gambling. If the social costs of EGM gambling include some fraction of police-recorded domestic violence as this study suggests, then this will increase estimates of social costs substantially. If such an analysis were to be carried out across the domains of gambling-related harm enumerated in the introduction, it seems likely that the true cost of gambling-related harm would outweigh estimates from the Productivity Commission (2010), for example, perhaps by an order of magnitude.

For gambling research, this study suggests that future research combining administrative data on EGM accessibility and complete administrative data on gambling-related harm provides a promising avenue by which to investigate the social impact of EGM gambling. Put simply, rather than trying to identify the location of problem gamblers (as much of the research in this thesis does), an alternative strategy of identifying the connection between increased EGM accessibility and increased incidence of specific gambling-related harms is likely to lead to improved knowledge of the connections between exposure to EGMs and public health impacts. While this approach is not entirely new (e.g. Grinols & Mustard 2006), the disciplinary focus on problem gambling (Cosgrave 2010; Miller et al. 2016; Reith 2007; Young 2013) has arguably resulted in an underdevelopment of this kind of research. Here gambling studies has much to learn from other cognate areas of public health research. For example, over 25 studies of the impact of alcohol-outlet density on specific public health outcomes have been published annually in recent years (Gmel et al. 2016). Gambling studies could benefit from this well-developed body of research, and from health geographic approaches more generally. For example, the research designs used to connect environmental exposures to health risk factors to health outcomes in the study of alcohol or tobacco could be replicated in the study of gambling. Similarly, environmental health interventions that have proven successful for these other dangerous commodities are likely to be successful – with appropriate adaptation – for gambling products.

Finally, this thesis has produced new evidence regarding the relationship between gambling expenditure and problem gambling at three spatial scales: the individual, the gambling venue and the jurisdiction. These three studies were consistent in finding that increased expenditure on EGMs was associated with an increased mass of gambling problems. This is perhaps the most important contribution of this thesis. In these cross-sectional studies, expenditure and harm rise and fall in concert. If this is the case, then several strategies for harm minimisation should be reconsidered. First, the production of low-risk limits identified on the basis of thresholds in J-shaped curves is likely to produce thresholds which are not meaningful from an epidemiological perspective, but rather reflect statistical artefacts. Second, all practices that serve to increase the demand for gambling will also be understood to increase the burden of gambling-related harm. Television advertising, for example, is likely to increase gambling consumption and thereby increase aggregate harms. A policy aimed at gambling harm minimisation would, therefore, restrict gambling advertising. Similar logic might apply to the licensing of EGM venues in local community spaces. Insofar as gambling accessibility stimulates the consumption of gambling by reducing the spatio-temporal barriers, it is also likely to increase harm. Indeed, once viewed through this heuristic, many routine gambling regulations – from the seeming small, like the denomination of currency accepted by EGMs, to the more significant, like the legalisation of new gambling products – may need to be reconsidered. If increased consumption necessarily leads to increased harm, then a public health approach to gambling regulation must favour policies that reduce consumption rather than increasing it.

This raises questions about the 'responsible gambling' model of gambling regulation. Specifically, the responsible gambling model suggests that gamblers are responsible for 'safe' (Blaszczynski et al. 2004) or 'healthy' (Korn & Shaffer 1999) gambling through practices such as 'seeking out information, setting limits on the amount of time and money he or she spends playing, making reasoned decisions, and controlling his or her own behaviour' (Reith, 2008, p. 150). Put differently, gambling and problem gambling must logically be constructed as being in opposition in responsible gambling discourse (Miller et al. 2016). As a policy goal, responsible gambling is attractive as it promises to resolve the contradictory interests of gambling industries which wish to see increased profits and a public who have an interest in minimising the burden of gambling-related harm. Responsible gambling provides a useful resolution of this contradiction because it implies that the harmfulness of gambling per dollar lost can be easily reduced. This thesis demonstrates that when it comes to EGMs, an easy resolution of this kind is likely to be impossible. Effective harm minimisation measures for EGMs will reduce the profitability of EGM proprietors. Consequently, harm minimisation for EGMs should be considered a political issue as much as a scientific-technical issue. Consequently, EGM policy will inevitably be the site of political conflict between the interests of gambling industries in extracting profits and the interests of the public in not being harmed by dangerous gambling products. As is the case in tobacco control, self-regulation or policies which are supported by the industry are unlikely to substantially reduce the burden of harm.

8.3 Limitations of the research

The studies compiled in this thesis are subject to many limitations. Specific limitations are detailed in each chapter, and need not be repeated here. However, two important caveats should be considered when evaluating the implications of these research findings. First, much of the research in this thesis is novel in the field of gambling research. Many of the research questions addressed in specific studies have not been previously raised. Consequently, there is a need for replication across contexts, datasets and methods to ensure that the findings presented in this thesis are generalizable, rather than being artefacts of specific datasets or idiosyncratic features of particular locales. Second, all of the research in this thesis is observational and concerned with detecting associations rather than making strong causal claims. This is appropriate because the identification of associations is important, both in its own right and as a precursor to testing for causal relations. Nevertheless, the findings of the research presented in this thesis should be considered with the requisite caution until independent replication and testing for causal relations can take place.

8.4 Recommendations for future work

All three strands of research require much future work be done. In the case of small area estimates, no external validation was able to be undertaken. External validation is necessary in order to evaluate the accuracy of spatial microsimulation derived estimates, and requires the existence of a third party 'gold standard' dataset against which estimates can be benchmarked. In the case of problem gambling prevalence in small areas, no such gold standard exists. A gold standard dataset requires the existence of a valid small area estimate with a large number of responses per spatial units. For some variables in some contexts, this may be sourced from administrative data, but no such data exist for gambling in Australia. However, in the Norwegian case, where EGMs consumption requires the use of a state-issued player-tracking card, such data may be obtainable. In other cases, census data may be used as a gold standard for validation. For example, in New Zealand, the quinquennial census asks about current cigarette smoking habits. This census question can be used to validate microsimulation-based estimates of smoking patterns. However, to the best of the authors' knowledge, no national census contains questions about gambling. Nor are prospects for any national census including questions about gambling good in future. Finally, a large enumeration survey (rather than sample survey) with full coverage of small areas could be used to obtain authoritative local estimates, although such an approach is likely to be very expensive.

Consequently, the development of creative strategies for externally validating these estimates – or the identification of a suitable benchmarking dataset – is needed to assess how accurate these small area estimates are in practice. Further, these methods have great potential to be applied in both regulatory and research contexts. First, if the promise of these methods for informing social impact assessment is to be realised, knowledge translation is required. In this case the dissemination of atlases of problem gambling prevalence to regulators, and perhaps the production of a set of user-friendly set of decision support tools for predicting problem gambling rates in the presence of changing EGM configurations, may be required in order for these new methods to realise their potential. Second, as argued above, the method for developing small area estimates described in this thesis should be applied to assist in answering research questions relating to the social and spatial distribution of the burden of gambling-related harm.

In the case of research linking expenditure and problem gambling, there is important further research to be done that takes a similar approach but does not use problem gambling as the outcome variable of interest. Specifically, there is important work to be done to estimate dose-response curves between losses on gambling products and specific gambling-related harms as the outcome variable (Rodgers et al. 2009). For example, little is known about the relationships between money lost on gambling and financial stress, or money lost on gambling and child neglect, or money lost on gambling and depression, or even money lost on gambling and all-cause mortality. Research investigating the predictors of specific gambling-related harms has tended to focus on problem gambling as the predictor variable instead of gambling consumption. This is a problem because, if both problem gambling and gambling-related harms result from gambling consumption, then these studies are comparing the result of two outcomes from gambling rather than an exposure and an outcome. Recent developments in conceptualising and measuring gambling-related harms may prove useful to such an investigation (Browne et al. 2016; Langham et al. 2016). Furthermore, several large, longitudinal datasets that track gambling behaviour among the same individuals over several years are now available (Abbott, Bellringer, et al. 2016; Billi et al. 2014; el-Guebaly et al. 2015; Romild et al. 2014; Williams et al. 2015). Analyses of these datasets will enable temporal sequencing to be established, a key requirement for causal inference. Alternatively, case control studies may be needed to provide evidence on the relationship between gambling consumption and relatively low-incidence, high-impact gambling-related harms such as attempted suicide or bankruptcy.

A spatial analytic approach can usefully provide a different method for answering the same research question at the community level. Future research could usefully focus on how the incidence of gambling-related harms in communities is predicted by the accessibility of EGMs. These studies might ask research questions about how, as accessibility to EGMs increases, the incidence of specific harms increases.

Methodologically, the latest research from cognate fields of spatial epidemiology are likely to provide a useful template for adaptation to the gambling context (for reviews of this literature, see Campbell et al. 2009; Gmel et al. 2016; Popova et al. 2009). In particular, studies with strong, quasi-experimental designs are likely to yield the strongest form of evidence.

8.5 Final conclusions

This thesis set out to investigate the impacts of EGMs, an entrenched social and public health issue, using health geographic approaches. In particular, it set out to develop small area estimates of problem gambling in order to inform regulation and research. This primary goal has been achieved, with the research in this thesis developing an empirically-calibrated set of prevalence estimates for spatial units of around 400 persons.

In doing so, this thesis has made several other contributions. It has applied existing geographic methods such as spatial interaction modelling to the domain of EGM gambling. More significantly, it has contributed to refining spatial microsimulation, a recently-developed geographical method, by demonstrating how microsimulation estimates can be improved by incorporating spatially-referenced ancillary data sources. This thesis has also made a significant contribution to the epidemiological understanding of the relationship of EGM gambling losses and problem gambling, investigating the association between EGM spending and harm at three spatial scales. Most importantly, this thesis has demonstrated that, contrary to the consensus understanding in the literature, the dose-response relationship between EGM losses and problem gambling risk is not J-shaped, a finding that has important implications for how EGMs should be regulated. Finally, this thesis has begun the work of applying health geographic methods to investigate the spatio-temporal relationships between EGM accessibility and gambling-related harms, an area of research that has been little developed in the academic literature.

The research developed in this thesis has contributed toward bringing knowledge of the geography of the impacts of EGMs closer to that of cognate public health issues such as alcohol and tobacco. In the main, it is hoped that the approaches developed in this thesis and the research findings will contribute to improving the regulation of EGMs and thereby reduce the incidence of gambling-related harms.

References

Abbott, M., 2005. Disabling the public interest: Gambling strategies and policies for Britain: A comment on Orford 2005. *Addiction*, 100(9), p.1233–1235.

Abbott, M., 2006. Do EGMs and problem gambling go together like a horse and carriage? *Gambling Research*, 18(1), p.7–38.

Abbott, M. et al., 2016. Gambling and Problem Gambling in Victoria, Australia: Changes over 5 years. *Journal of Gambling Studies*, 32(1), p.1–32.

Abbott, M. et al., 2016. *New Zealand National Gambling Study: Wave 3*, Auckland: Ministry of Health. Available at: https://niphmhr.aut.ac.nz/__data/assets/pdf_file/0010/62668/Final-Report-National-Gambling-Study-Wave-3-19.08.16.pdf archived at http://www.webcitation.org/6px7HPv2D [Accessed April 24, 2017].

Abbott, M.W., Volberg, R.A. & Williams, M.M., 1999. *Seven years on: A follow-up study of frequent and problem gamblers living in the community*, Wellington, New Zealand: Department of Internal Affairs.

Afifi, T.O. et al., 2010. The relationship of gambling to intimate partner violence and child maltreatment in a nationally representative sample. *Journal of Psychiatric Research*, 44(5), p.331–337.

Allen Consulting Group, 2012. *Social and Economic Impact Study of Gambling in Tasmania, Volume 1: Gambling Industry Trends and Economic Impacts*, Hobart: Tasmanian Government Department of Treasury and Finance. Available at: http://www.treasury.tas.gov.au/Documents/Volume1secondgamblingSEIS.PDF archived at http://www.stors.tas.gov.au/au-7-0095-06779 [Accessed August 22, 2012].

Al-Yaman, F., Van Doeland, M. & Wallis, M., 2006. *Family Violence Among Aboriginal and Torres Strait Islander Peoples*, Canberra: Australian Institute of Health and Welfare. Available at: http://www.aihw.gov.au/publication-detail/?id=6442467912 archived at http://www.webcitation.org/6d6BytMUn [Accessed November 17, 2015].

American Gaming Association, 2012. *The State of the States: The AGA Survey of Casino Entertainment, 2012 ed.*, Washington: American Gaming Association. Available at: https://www.americangaming.org/research/reports/state-states-12 archived at http://www.webcitation.org/6q2KhdQ78 [Accessed April 20, 2017].

American Psychiatric Association, 2013. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, Arlington, Va: American Psychiatric Association.

Anderson, B., 2007. *Creating Small Area Income Estimates for England: Spatial microsimulation modelling*, Ipswitch, UK: University of Essex. Available at: http://opendepot.org/166 archived at http://www.webcitation.org/6mPHIxL5g [Accessed December 1, 2016].

Australian Bureau of Statistics, 2008. *An Introduction to Socio-Economic Indexes for Areas (SEIFA), 2006*, Canberra: Australian Bureau of Statistics. Available at: http://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0/ archived at http://www.webcitation.org/67tGG7kJi [Accessed May 24, 2012].

Australian Bureau of Statistics, 2016. *Australian System of National Accounts, 2014-15*, Canberra: Australian Bureau of Statistics. Available at: http://www.abs.gov.au/ausstats/abs@.nsf/mf/5204.0 archived at http://www.webcitation.org/6q2KwqEwr [Accessed February 10, 2016].

Australian Bureau of Statistics, 2011. *Regional Population Growth, Australia, 2009-10*, Canberra: Australian Bureau of Statistics. Available at: http://www.abs.gov.au/ausstats/abs@.nsf/mf/3218.0 [Accessed January 1, 2018].

Australian Bureau of Statistics, 2012. *Personal Safety Survey - Expanded CURF, RADL. Findings based on use of ABS CURF data.*, Canberra: Australian Bureau of Statistics.

Australian Bureau of Statistics, 2015. *Regional Population Growth, Australia, 2013-14*, Canberra: Australian Bureau of Statistics. Available at: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3218.02013-14?OpenDocument [Accessed November 16, 2015].

Australian Bureau of Statistics, 2013. *Technical Paper: Socio-Economic Indexes for Areas (SEIFA), Australia*, Canberra. Available at: http://www.abs.gov.au/ausstats/abs@.nsf/mf/2033.0.55.001 archived at http://www.webcitation.org/6mQ75xZ0Q [Accessed December 1, 2016].

Australian Competition and Consumer Commission, 2008. *Report of the ACCC Inquiry into the Competitiveness of Retail Prices for Standard Groceries*, Canberra: Commonwealth of Australia.

Australian Institute for Gambling Research, 1999. *Australian Gambling Comparative History and Analysis*, Melbourne: Victorian Casino and Gaming Authority. Available at:

https://assets.justice.vic.gov.au/vcglr/resources/bb81f943-d854-40de-8bab-b09d8bbd610f/australiangamblingcomparativehistory.pdf archived at http://www.webcitation.org/6cNpqq7Xn [Accessed October 19, 2015].

Australian Psychological Society, 2015. *Stress and Wellbeing: How Australians are Coping with Life*, Melbourne. Available at: http://www.psychology.org.au/Assets/Files/PW15-SR.pdf archived at http://www.webcitation.org/6d4TELRHo [Accessed November 15, 2015].

Babor, T. et al., 2010. *Alcohol: No Ordinary Commodity: Research and Public Policy* 2nd ed., Oxford: Oxford University Press.

Barratt, M.J. et al., 2014. Gaming machine density is correlated with rates of help-seeking for problem gambling: a local area analysis in Victoria, Australia. *Journal of Gambling Issues*, 29, p.1–21.

Barry, T., 1999. The National Visitor Survey. *Journal of the Bureau of Tourism Research*, 1(1), p.1–8.

Batty, M. & Sikdar, P.K., 1984. Proximate aggregation-estimation of spatial interaction models. *Environment and Planning A*, 16(4), p.467 – 486.

Batty, M. & Sikdar, P.K., 1982. Spatial aggregation in gravity models: 4. Generalisations and large-scale applications. *Environment and Planning A*, 14(6), p.795 – 822.

Beyer, K., Wallis, A.B. & Hamberger, L.K., 2015. Neighborhood environment and intimate partner violence: A systematic review. *Trauma, Violence, & Abuse*, 16(1), p.16–47.

Billi, R. et al., 2014. *The Victorian Gambling Study: A Longitudinal Study of Gambling and Health in Victoria 2008-2012*, North Melbourne: Victorian Responsible Gambling Foundation. Available at: http://www.responsiblegambling.vic.gov.au/about-us/what-we-do/foundation-research-program/research-programs-and-activities/victorian-gambling-study archived at http://www.webcitation.org/6b1SfoUPD [Accessed August 24, 2015].

Bivand, R.S., Pebesma, E.J. & Gómez-Rubio, V., 2008. *Applied Spatial Data Analysis with R*, New York: Springer.

Blaszczynski, A., Ladouceur, R. & Shaffer, H.J., 2004. A science-based framework for responsible gambling: The Reno model. *Journal of Gambling Studies*, 20(3), p.301–317.

Blaszczynski, A. & Nower, L., 2002. A pathways model of problem and pathological gambling. *Addiction*, 97(5), p.487–499.

Borenstein, M. et al., 2009. *Introduction to Meta-Analysis*, Chichester, U.K: Wiley.

Browne, M. et al., 2016. *Assessing gambling-related harm in Victoria: a public health perspective*, Melbourne: Victorian Responsible Gambling Foundation. Available at: https://www.responsiblegambling.vic.gov.au/__data/assets/pdf_file/0007/28465/Browne_assessing_gambling-related_harm_in_Vic_Apr_2016-REPLACEMENT2.pdf archived at http://www.webcitation.org/6pEx9apZZ [Accessed November 29, 2016].

Bruun, K. et al., 1975. *Alcohol Control Policies in Public Health Perspective*, Helsinki: The Finnish Foundation for Alcohol Studies.

Bunkle, P. & Lepper, J., 2004. What do we know about gambling in New Zealand? *Social Policy Journal of New Zealand*, 21(Mar), p.178–201.

van Buuren, S. & Groothuis-Oudshoorn, K., 2011. MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), p.1–67.

Campbell, C.A. et al., 2009. The effectiveness of limiting alcohol outlet density as a means of reducing excessive alcohol consumption and alcohol-related harms. *American Journal of Preventive Medicine*, 37(6), p.556–569.

Campbell, J.C., 2002. Health consequences of intimate partner violence. *The Lancet*, 359(9314), p.1331–1336.

Canadian Partnership for Responsible Gambling, 2015. *Canadian Gambling Digest, 2013-2014*, Toronto: Responsible Gambling Council. Available at: http://www.responsiblegambling.org/docs/default-source/default-document-library/cprg_canadian-gambling-digest_2013-14.pdf?sfvrsn=2 archived at http://www.webcitation.org/6p1YFiWuH [Accessed January 27, 2016].

Castrén, S. et al., 2013. Factors associated with disordered gambling in Finland. *Substance Abuse Treatment, Prevention, and Policy*, 8(24), p.1–10.

Cataife, G., 2014. Small area estimation of obesity prevalence and dietary patterns: A model applied to Rio de Janeiro city, Brazil. *Health & Place*, 26(Mar), p.47–52.

Chambers, K., 2011. *Gambling for Profit: Lotteries, Gaming Machines, and Casinos in Cross-national Focus*, Toronto: University of Toronto Press.

Chokshi, D.A., El-Sayed, A.M. & Stine, N.W., 2015. J-Shaped curves and public health. *JAMA*, 314(13), p.1339–1340.

Clarke, D. et al., 2010. An exploratory study of problem gambling on casino versus non-casino electronic gaming machines. *International Journal of Mental Health and Addiction*, 10(1), p.107–121.

Cleveland, W., Grosse, E. & Shyu, W., 1992. Local regression models. In J. Chambers & T. Hastie, eds. *Statistical Models in S*. Pacific Grove, California: Wadsworth & Brooks/Cole, pp. 309–376.

Collins, D. & Lapsley, H., 2008. *The Costs of Tobacco, Alcohol and Illicit Drug Abuse to Australian Society in 2004/05*, Canberra: Australian Government Department of Health and Ageing. Available at: http://www.health.gov.au/internet/drugstrategy/publishing.nsf/Content/mono64 archived at http://www.webcitation.org/6pFFOz3zj [Accessed September 30, 2016].

Conway, M., 2015. Vulnerability modeling of casinos in the United States: A case study of Philadelphia. *Applied Geography*, 63(2015), p.21–32.

Cosgrave, J.F., 2010. Embedded addiction: The social production of gambling knowledge and the development of gambling markets. *The Canadian Journal of Sociology*, 35(1), p.113–134.

Cullinan, J. et al., 2012. Accessibility and screening uptake rates for gestational diabetes mellitus in Ireland. *Health & Place*, 18(2), p.339–348.

Cunradi, C.B. et al., 2011. Alcohol outlets, neighborhood characteristics, and intimate partner violence: Ecological analysis of a Californian city. *Journal of Urban Health*, 88(2), p.191–200.

Currie, S.R. et al., 2008. Replication of low-risk gambling limits using Canadian provincial gambling prevalence data. *Journal of Gambling Studies*, 24(3), p.321–335.

Currie, S.R. et al., 2006. Risk of harm among gamblers in the general population as a function of level of participation in gambling activities. *Addiction*, 101(4), p.570–580.

Currie, S.R., Hodgins, D.C. & Casey, D.M., 2013. Validity of the Problem Gambling Severity Index Interpretive Categories. *Journal of Gambling Studies*, 29(2), p.311–327.

Delfabbro, P. & King, D., 2017. Gambling is not a capitalist conspiracy: a critical commentary of literature on the 'industry state gambling complex'. *International Gambling Studies*, online publication prior to print, p.1–15.

DeVerteuil, G., 2015. Conceptualizing violence for health and medical geography. *Social Science & Medicine*, 133(2015), p.216–222.

Devries, K.M. et al., 2013. The global prevalence of intimate partner violence against women. *Science*, 340(6140), p.1527–1528.

Doran, B., Marshall, D. & McMillen, J., 2007. A GIS-based investigation of gaming venue catchments. *Transactions in GIS*, 11(4), p.575–595.

Doran, B. & Young, M., 2010. Predicting the spatial distribution of gambling vulnerability: An application of gravity modeling using ABS Mesh Blocks. *Applied Geography*, 30(1), p.141–152.

Doughney, J., 2009. Lies, damned lies and 'problem gambling' prevalence rates: The example of Victoria, Australia. *Journal of Business Systems, Governance and Ethics*, 2(1), p.41–54.

Dowling, N. et al., 2016. Problem gambling and intimate partner violence: A systematic review and meta-analysis. *Trauma, Violence, & Abuse*, 17(1), p.43–61.

Dowling, N.A. et al., 2016. National estimates of Australian gambling prevalence: findings from a dual-frame omnibus survey. *Addiction*, 111(3), p.420–435.

Eadington, W.R., 1998. Casino gaming - origins, trends and impacts. In K. J. Meyer-Arendt & R. Hartmann, eds. *Casino Gambling in America: Origins, trends and impacts*. New York: Cognizant Communication Corporation, pp. 3–15.

ESRI, 2010. *ArcGIS Desktop*, Redlands, CA: Environmental Systems Research Institute.

Evans, W.N. & Topoleski, J.H., 2002. *The Social and Economic Impact of Native American Casinos*, National Bureau of Economic Research. Available at: http://www.nber.org/papers/w9198 archived at http://www.webcitation.org/6q2MdM0op [Accessed January 13, 2013].

Felson, R.B. et al., 2002. Reasons for reporting and not reporting domestic violence to the police. *Criminology*, 40(3), p.617–648.

Ferris, J. & Wynne, H., 2001a. *The Canadian Problem Gambling Index: Final Report*, Canadian Centre on Substance Abuse. Available at:

http://www.cclat.ca/2003%20and%20earlier%20CCSA%20Documents/ccsa-008805-2001.pdf archived at http://www.webcitation.org/66rzvX4ZE [Accessed November 10, 2011].

Ferris, J. & Wynne, H., 2001b. *The Canadian Problem Gambling Index: User Manual*, Ottawa: Canadian Centre on Substance Abuse. Available at: http://www.ccsa.ca/2003%20and%20earlier%20CCSA%20Documents/ccsa-009381-2001.pdf archived at http://www.webcitation.org/67tGlKq7P [Accessed November 10, 2011].

Ferris, J., Wynne, H. & Single, E., 2000. *The Canadian Problem Gambling Index 2000 [Canada]: Canadian National Validation Survey [computer file]*, Toronto: Gambling Research Exchange Ontario [distributor]. Available at: http://hdl.handle.net/10864/10724 UNF:5:P3hj0yVXmrUpfopfFjZYDw== [Accessed November 10, 2011].

Fleming, D.A. & Measham, T.G., 2015. Income inequality across Australian regions during the mining boom: 2001–11. *Australian Geographer*, 46(2), p.203–216.

Fotheringham, S. & O'Kelly, M.E., 1989. *Spatial Interaction Models: Formulations and Applications*, Dordrecht ; Boston: Kluwer Academic Publishers.

Francis, L., Livingstone, C. & Rintoul, A., 2017. Analysis of EGM licensing decisions by the gambling regulator, Victoria, Australia. *International Gambling Studies*, 17(1), p.65–86.

Freisthler, B. & Weiss, R.E., 2008. Using Bayesian space-time models to understand the substance use environment and risk for being referred to child protective services. *Substance Use & Misuse*, 43(2), p.239–251.

Gainsbury, S.M. et al., 2014. The prevalence and determinants of problem gambling in Australia: Assessing the impact of interactive gambling and new technologies. *Psychology of Addictive Behaviors*, 28(3), p.769–779.

Gautschi, D.A., 1981. Specification of patronage models for retail center choice. *Journal of Marketing Research*, 18(2), p.162–174.

Gelman, A. & Stern, H., 2006. The Difference Between 'Significant' and 'Not Significant' is not Itself Statistically Significant. *The American Statistician*, 60(4), p.328–331.

Gerstein, D. et al., 1999. *Gambling Impact and Behavior Study: Report to the National Gambling Impact Study Commission*, Chicago: National Opinion and Research Centre, the University of Chicago.

Gmel, G., Holmes, J. & Studer, J., 2016. Are alcohol outlet densities strongly associated with alcohol-related outcomes? A critical review of recent evidence. *Drug and Alcohol Review*, 35(1), p.40–54.

Gracia, E. et al., 2014. Exploring neighborhood influences on small-area variations in intimate partner violence risk: A Bayesian random-effects modeling approach. *International Journal of Environmental Research and Public Health*, 11(1), p.866–882.

Gracia, E. et al., 2015. The spatial epidemiology of intimate partner violence: Do neighborhoods matter? *American Journal of Epidemiology*, 182(1), p.58–66.

Gracia, E., 2004. Unreported cases of domestic violence against women: towards an epidemiology of social silence, tolerance, and inhibition. *Journal of Epidemiology and Community Health*, 58(7), p.536–537.

Grinols, E.L. & Mustard, D.B., 2006. Casinos, crime, and community costs. *Review of Economics and Statistics*, 88(1), p.28–45.

Grun, L. & McKeigue, P., 2000. Prevalence of excessive gambling before and after introduction of a national lottery in the United Kingdom: another example of the single distribution theory. *Addiction*, 95(6), p.959–966.

el-Guebaly, N. et al., 2015. *The Leisure, Lifestyle and Lifecycle Project (LLLP): A Longitudinal Study of Gambling in Alberta*, Alberta Gambling Research Institute. Available at: http://hdl.handle.net/1880/50377 archived at http://www.webcitation.org/6b1SNrNPU [Accessed August 24, 2015].

Haining, R., 2003. *Spatial Data Analysis: Theory and Practice*, Cambridge: Cambridge University Press.

Hair, J.F. et al., 2009. *Multivariate Data Analysis* 7 edition., Upper Saddle River, NJ: Pearson.

Hamby, S., 2014. Intimate partner and sexual violence research: Scientific progress, scientific challenges, and gender. *Trauma, Violence, & Abuse*, 15(3), p.149–158.

Hancock, L., Hao, Z., 2016. Gambling regulatory regimes and the framing of “responsible gambling” by transnational casino corporations: Asia-Pacific regimes in comparative perspective. *Asia Pacific Journal of Public Administration* 38, 139–153.

Hansen, M. & Rossow, I., 2008. Adolescent gambling and problem gambling: Does the total consumption model apply? *Journal of Gambling Studies*, 24(2), p.135–149.

Hansen, W., 1959. How Accessibility Shapes Land Use. *Journal of the American Planning Association*, 25(2), p.73–76.

Heise, L.L., 1998. Violence against women: An integrated, ecological framework. *Violence Against Women*, 4(3), p.262–290.

Heise, L.L. & Kotsadam, A., 2015. Cross-national and multilevel correlates of partner violence: an analysis of data from population-based surveys. *The Lancet Global Health*, 3(6), p.e332–e340.

Herd, D., 1985. Migration, Cultural Transformation and the Rise of Black Liver Cirrhosis Mortality. *British Journal of Addiction* 80, 397–410.

Higgins, J.P.T. et al., 2003. Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), p.557–560.

Higgins, J.P.T. & Thompson, S.G., 2004. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, 23(11), p.1663–1682.

Higgins, J.P.T. & Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), p.1539–1558.

Hirschberg, J.G. & Lye, J.N., 2010. The indirect impacts of smoking bans in gaming venues. In D. Slottje & R. Tchernis, eds. *Current Issues in Health Economics*. Bingley , UK: Emerald Group Publishing Limited, pp. 243–259.

Huff, D.L., 1963. A probabilistic analysis of shopping center trade areas. *Land Economics*, 39(1), p.81–90.

Huff, D.L., 1964. Defining and estimating a trading area. *The Journal of Marketing*, 28(3), p.34–38.

Huff, D.L. & McCallum, B., 2008. *Calibrating the Huff model using ArcGIS business analyst*, Redlands, California: ESRI.

Hurvich, C.M., Simonoff, J.S. & Tsai, C.-L., 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2), p.271–293.

Jackson, A.C. et al., 2010. Using the CPGI to Determine Problem Gambling Prevalence in Australia: Measurement Issues. *International Journal of Mental Health and Addiction*, 8(4), p.570–582.

Jellinek, E.M., 1952. Phases of alcohol addiction. *Q. J. Stud. Alcohol* 13, 673–684.

Jerrett, M., Gale, S. & Kontgis, C., 2009. An Environmental Health Geography of Risk. In T. Brown, S. McLafferty, & G. Moon, eds. *A Companion to Health and Medical Geography*. London: Wiley-Blackwell, pp. 418–445.

Jewkes, R., 2002. Intimate partner violence: causes and prevention. *The Lancet*, 359(9315), p.1423–1429.

Kavroudakis, D., 2015. sms: An R Package for the Construction of Microdata for Geographical Analysis. *Journal of Statistical Software*, 68(2), p.1–23.

Knorr-Held, L., 2000. Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*, 19(17–18), p.2555–2567.

Korman, L.M. et al., 2008. Problem Gambling and Intimate Partner Violence. *Journal of Gambling Studies*, 24(1), p.13–23.

Korn, D.A. & Shaffer, H.J., 1999. Gambling and the Health of the Public: Adopting a Public Health Perspective. *Journal of Gambling Studies*, 15(4), p.289–365.

KPMG Consulting, 2000. *Report of the 1999 Longitudinal Community Impact Study*, Melbourne: Victorian Casino and Gaming Authority.

Krantz, G., 2002. Violence against women: a global public health issue! *Journal of Epidemiology and Community Health*, 56(4), p.242–243.

Lal, A. & Siahpush, M., 2008. The effect of smoke-free policies on electronic gaming machine expenditure in Victoria, Australia. *Journal of Epidemiology & Community Health*, 62(1), p.11–15.

Lambert, P.C. et al., 2005. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine*, 24(15), p.2401–2428.

Langham, E. et al., 2016. Understanding gambling related harm: a proposed definition, conceptual framework, and taxonomy of harms. *BMC Public Health*, 16:80, p.1–23.

Lattimore, R. & Phillips, R., 2000. *Gambling Inquiry Survey, 1999 [computer file]*, Canberra: Australian Data Archive, The Australian National University.

Ledermann, S., 1956. Alcool, alcoolisme, alcoolisation. Presses univ. de France, Paris.

Lee, D., Rushworth, A. & Napier, G., 2015. *CARBayesST: An R Package for Spatio-temporal Areal Unit Modelling with Conditional Autoregressive Priors*, Vienna, Austria: The Comprehensive R Archive Network.

Lesieur, H.R. & Blume, S.B., 1987. The South Oaks gambling screen (SOGS): a new instrument for the identification of problem gamblers. *American Journal of Psychiatry*, 144(9), p.1184–1188.

Lim, S.S. et al., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), p.2224–2260.

Little, R.J. et al., 2012. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*, 367(14), p.1355–1360.

Livingstone, C., Rintoul, A. & Francis, L., 2014. What is the evidence for harm minimisation measures in gambling venues? *Evidence Base*, 2014(2), p.1–24.

Livingstone, C. & Woolley, R., 2007. Risky business: A few provocations on the regulation of electronic gaming machines. *International Gambling Studies*, 7(3), p.361–376.

Lovelace, R. & Dumont, M., 2016. *Spatial Microsimulation with R*, Boca Raton, FL: CRC Press.

Lund, I., 2006. Gambling and problem gambling in Norway: What part does the gambling machine play? *Addiction Research & Theory*, 14(5), p.475–491.

Lund, I., 2008. The population mean and the proportion of frequent gamblers: Is the theory of total consumption valid for gambling? *Journal of Gambling Studies*, 24(2), p.247–256.

Macdonald, L., Olsen, J.R., Shortt, N.K., Ellaway, A., 2018. Do ‘environmental bads’ such as alcohol, fast food, tobacco, and gambling outlets cluster and co-locate in more deprived areas in Glasgow City, Scotland? *Health & Place* 51, 224–231.

MacKinnon, J.G. & White, H., 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3), p.305–325.

MacLaren, V.V., 2016. Video lottery is the most harmful form of gambling in Canada. *Journal of Gambling Studies*, 32(2), p.459–485.

Markham, F., 2015. The open accessibility and remoteness index for Australia: Technical report. *figshare*. Available at: http://doi.org/10.6084/m9.figshare.1574190 archived at http://www.webcitation.org/6q2NtQjNG [Accessed October 15, 2015].

Markham, F., Doran, B. & Young, M., 2014a. Estimating gambling venue catchments for impact assessment using a calibrated gravity model. *International Journal of Geographical Information Science*, 28(2), p.326–342.

Markham, F., Doran, B. & Young, M., 2014b. The spatial extents of casino catchments in Australia. *Growth and Change*, 45(1), p.60–78.

Markham, F., Doran, B., Young, M., 2016. The relationship between electronic gaming machine accessibility and police-recorded domestic violence: A spatio-temporal analysis of 654 postcodes in Victoria, Australia, 2005–2014. *Social Science & Medicine* 162(2016), p.106–114.

Markham, F., Young, M., 2015. "Big Gambling": The rise of the global industry-state gambling complex. *Addiction Research & Theory,* 23(2015), p.1–4.

Markham, F. & Young, M., 2016. Commentary on Dowling et al. (2016): Is it time to stop conducting problem gambling prevalence studies? *Addiction*, 111(3), p.436–437.

Markham, F., Young, M. & Doran, B., 2014. Gambling expenditure predicts harm: evidence from a venue-level study. *Addiction*, 109(9), p.1509–1516.

Markham, F., Young, M. & Doran, B., 2016. The relationship between player losses and gambling-related harm: evidence from nationally representative cross-sectional surveys in four countries. *Addiction*, 111(2), p.320–330.

Markham, F., Young, M., Doran, B., 2017a. Improving spatial microsimulation estimates of health outcomes by including geographic indicators of health behaviour: The example of problem gambling. *Health & Place* 46(2017), p.29–36.

Markham, F. et al., 2017b. A meta-regression analysis of 41 Australian problem gambling prevalence estimates and their relationship to total spending on electronic gaming machines. *BMC Public Health,* 17(495), p.1–12.

Marmot, M., 2005. Social determinants of health inequalities. *The Lancet,* 365, p.1099–1104.

Marshall, D., 1999. Adelaide's pokie geography: distribution of, and expenditure on, electronic gaming machines in Adelaide. *South Australian Geographical Journal*, 98(1999), p.19–29.

Marshall, D., 2005. The gambling environment and gambler behaviour: Evidence from Richmond-Tweed, Australia. *International Gambling Studies*, 5(1), p.63–83.

Marshall, D. & Baker, R.G.V., 2001a. Clubs, spades, diamonds and disadvantage: The geography of electronic gaming machines in Melbourne. *Australian Geographical Studies*, 39(1), p.17–33.

Marshall, D. & Baker, R.G.V., 2000. No chance: electronic gaming machines and the disadvantaged in Melbourne. *Geodate*, 13(4), p.5–8.

Marshall, D. & Baker, R.G.V., 2002. The evolving market structures of gambling: Case studies modelling the socioeconomic assignment of gaming machines in Melbourne and Sydney, Australia. *Journal of Gambling Studies*, 18(3), p.273–291.

Marshall, D. & Baker, R.G.V., 2001b. Unfair odds? Factors influencing the distribution of electronic gaming machines in Melbourne. *Urban Policy and Research*, 19(1), p.77–92.

McCarthy, G., 1999. The 1997 South Australian election: history never repeats. *Legislative Studies*, 13(2), p.61–77.

McCready, J. & Adlaf, E., 2006. *Performance and Enhancement of the Canadian Problem Gambling Index: Report and Recommendations*, Canadian Consortium for Gambling Research. Available at: http://www.ccgr.ca/sites/default/files/Performance-and-Enhancement-of-the-Canadian-Problem-Gambling-Index-CPGI-Report-and-Recommendations.pdf archived at http://www.webcitation.org/6p1Ypb1Cg [Accessed February 7, 2016].

McMillen, J. & Doran, B., 2006. Problem gambling and gaming machine density: Socio-spatial analysis of three Victorian localities. *International Gambling Studies*, 6(1), p.5–29.

Midanik, L.T. et al., 1996. Risk functions for alcohol-related problems in a 1988 US national sample. *Addiction*, 91(10), p.1427–1437.

Miller, H.E. et al., 2016. Surveillance, responsibility and control: an analysis of government and industry discourses about 'problem' and 'responsible' gambling. *Addiction Research & Theory*, 24(2), p.163–176.

Mouzos, J. & Makkai, T., 2004. *Women's Experiences of Male Violence: Findings from the Australian Component of the International Violence Against Women Survey (IVAWS)*, Canberra: Australian Institute of Criminology.

Muelleman, R.L. et al., 2002. Problem gambling in the partner of the emergency department patient as a risk factor for intimate partner violence. *The Journal of Emergency Medicine*, 23(3), p.307–312.

Nagle, N.N. et al., 2014. Dasymetric modeling and uncertainty. *Annals of the Association of American Geographers*, 104(1), p.80–95.

Nakaya, T. et al., 2007. Combining microsimulation and spatial interaction models for retail location analysis. *Journal of Geographical Systems*, 9(4), p.345–369.

National Institute for Health and Welfare, 2013. *Gambling Survey 2011 [computer file]. FSD2722, version 2.0.*, Helsinki: Finnish Social Science Data Archive.

Neal, P., Delfabbro, P.H. & O'Neil, M., 2005. *Problem Gambling and Harm: Towards a National Definition*, Melbourne: Victorian Department of Justice. Available at: https://www.adelaide.edu.au/saces/docs/problemgamblingandharmtowardnationaldefinition.pdf archived at http://www.webcitation.org/6pz7uT2b4 [Accessed April 20, 2017].

Newman, O., 1972. *Gambling: Hazard and Reward*, London: Athlone Press. Available at: http://prism.ucalgary.ca//handle/1880/41340 [Accessed February 10, 2016].

O'Donnell, C.J., Smith, A. & Madison, J.R., 2002. Using demographic risk factors to explain variations in the incidence of violence against women. *Journal of Interpersonal Violence*, 17(12), p.1239–1262.

Office of Economic and Statistical Research, 2012. *Australian Gambling Statistics, 28th Edition*, Brisbane, Qld: Queensland Treasury and Trade. Available at: http://www.oesr.qld.gov.au/products/publications/aus-gambling-stats/index.php archived at http://www.webcitation.org/6q2OKdEsi [Accessed August 6, 2013].

Openshaw, S., 1977. Optimal zoning systems for spatial interaction models. *Environment and Planning A*, 9(2), p.169–184.

Openshaw, S., 1984. *The Modifiable Areal Unit Problem*, Norwich: Geo Books.

Orford, J., Wardle, H. & Griffiths, M., 2013. What proportion of gambling is problem gambling? Estimates from the 2010 British Gambling Prevalence Survey. *International Gambling Studies*, 13(1), p.4–18.

Orwin, R.G. & Vevea, J.L., 2009. Evaluating coding decisions. In H. Cooper, L. V. Hedges, & J. C. Valentine, eds. *Handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation, pp. 177–203.

Pearce, J. et al., 2008. A national study of neighborhood access to gambling opportunities and individual gambling behaviour. *Journal of Epidemiology and Community Health*, 62(10), p.862–868.

Pennay, A., Holmes, J., Törrönen, J., Livingston, M., Kraus, L., Room, R., 2018. Researching the decline in adolescent drinking: The need for a global and generational approach. *Drug and Alcohol Review* 37, p.S115–S119.

Peters, E. et al., 2015. Relationship of gambling with tobacco, alcohol, and illicit drug use among adolescents in the USA: Review of the literature 2000–2014. *The American Journal on Addictions* 24(3), p.206–216.

Phillips, S.P., 2011. Including gender in public health research. *Public Health Reports*, 126, p.16–21.

Pilver, C.E., Potenza, M.N., 2013. Increased incidence of cardiovascular conditions among older adults with pathological gambling features in a prospective study. *Journal of Addiction Medicine*, 7(6), p.387–393.

Pinchevsky, G.M. & Wright, E.M., 2012. The impact of neighbourhoods on intimate partner violence and victimization. *Trauma, Violence, & Abuse*, 13(2), p.112–132.

Plummer, M., 2015. *JAGS Version 4.0.0 User Manual*, Lyon, France: International Agency for Research on Cancer, World Health Organisation. Available at: http://sourceforge.net/projects/mcmc-jags/files/Manuals/4.x/ archived at http://www.webcitation.org/6p1YQLhhG [Accessed December 12, 2016].

Popova, S. et al., 2009. Hours and days of sale and density of alcohol outlets: Impacts on alcohol consumption and damage: A systematic review. *Alcohol and Alcoholism*, 44(5), p.500–516.

Productivity Commission, 1999. *Australia's Gambling Industries*, Canberra: Productivity Commission.

Productivity Commission, 2010. *Gambling*, Canberra: Productivity Commission.

PSMA Australia, 2010. *Product Description: G-NAF*, Canberra: PSMA Australia. Available at: http://www.psma.com.au/psma/wp-content/uploads/G-NAF-Product-Description.pdf archived at http://www.webcitation.org/6OK6xR5rP [Accessed March 24, 2014].

Queensland Government Statistician's Office, 2016. *Australian Gambling Statistics, 1989–90 to 2014–15, 32nd edition*, Brisbane: Queensland Treasury. Available at: http://www.qgso.qld.gov.au/products/reports/aus-gambling-stats/aus-gambling-stats-32nd-edn.pdf archived at http://www.webcitation.org/6p1XPNg4C [Accessed January 5, 2017].

R Core Team, 2015. *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.R-project.org.

R Development Core Team, 2012. *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing.

Reece, W.S., 2010. Casinos, hotels, and crime. *Contemporary Economic Policy*, 28(2), p.145–161.

Rehm, J., Kehoe, T., Gmel, Gerrit, Stinson, F., Grant, B., Gmel, Gerhard, 2010. Statistical modeling of volume of alcohol exposure for epidemiological studies of population health: the US example. *Population Health Metrics* 8(3).

Reith, G., 2008. Editorial: Reflections on responsibility. *Journal of Gambling Issues*, 22(December), p.149–155.

Reith, G., 2007. Gambling and the contradictions of consumption: A genealogy of the 'pathological' subject. *American Behavioral Scientist*, 51(1), p.33–55.

Rintoul, A.C. et al., 2013. Modelling vulnerability to gambling related harm: How disadvantage predicts gambling losses. *Addiction Research & Theory*, 21(4), p.329–338.

Robitaille, E. & Herjean, P., 2008. An analysis of the accessibility of video lottery terminals: the case of Montreal. *International Journal of Health Geographics*, 7(2), p.1–15.

Rockloff, M.J., 2012. Validation of the consumption screen for problem gambling (CSPG). *Journal of Gambling Studies*, 28(2), p.207–216.

Rockloff, M.J., Greer, N., Fay, C., 2011. The Social Contagion of Gambling: How Venue Size Contributes to Player Losses. *J Gambl Stud* 27, 487–497.

Rodgers, B. et al., 2015. *Gambling Expenditure in the ACT (2009): By Level of Problem Gambling, Type of Activity and Socioeconomic and Demographic Characteristics*, Canberra: Centre for Gambling Research, The Australian National University. Available at: http://www.gamblingandracing.act.gov.au/__data/assets/pdf_file/0010/982774/2014-Gambling-Expenditure.pdf archived at http://www.webcitation.org/6pz9VPWzv [Accessed January 3, 2017].

Rodgers, B., Caldwell, T. & Butterworth, P., 2009. Measuring gambling participation. *Addiction*, 104(7), p.1065–1069.

Romild, U., Volberg, R. & Abbott, M., 2014. The Swedish longitudinal gambling study (swelogs): Design and methods of the epidemiological track. *International Journal of Methods in Psychiatric Research*, 23(3), p.372–386.

Ronksley, P.E. et al., 2011. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. *BMJ*, 342(d671), p.1–13.

Rönnberg, S. et al., 1999. *Gambling and problem gambling in Sweden*, National Institute of Public Health. Available at: http://www.jogoremoto.pt/docs/extra/DeSt8m.pdf archived at http://www.webcitation.org/6OSYULsFI [Accessed March 30, 2014].

Room, R., Livingston, M., 2017. The Distribution of Customary Behavior in a Population: The Total Consumption Model and Alcohol Policy. *Sociological Perspectives* 60, p.10–22.

Room, R., Turner, N.E. & Ialomiteanu, A., 1999. Community effects of the opening of the Niagara casino. *Addiction*, 94(10), p.1449–1466.

Rose, G. & Day, S., 1990. The population mean predicts the number of deviant individuals. *BMJ: British Medical Journal*, 301(6759), p.1031–1034.

Rosenberg, M., 2014. Health geography I: Social justice, idealist theory, health and health care. *Progress in Human Geography*, 38(3), p.466–475.

Roy Morgan Research, 1999. *Sixth Survey of Community Gambling Patterns and Perceptions*, Melbourne: Victorian Casino and Gaming Authority. Available at: https://assets.justice.vic.gov.au/vcglr/resources/9412cce8-63fe-4f5e-82e0-fe5c0951b62a/6thsurveycommunitygamblingpattersnandperceptions.pdf archived at http://www.webcitation.org/6pqrAjoG1 [Accessed October 1, 2016].

Rush, B., Veldhuizen, S. & Adlaf, E., 2007. Mapping the prevalence of problem gambling and its association with treatment accessibility and proximity to gambling venues. *Journal of Gambling Issues*, 20(2007), p.193–214.

Sassen, M., Kraus, L. & Bühringer, G., 2011. Differences in pathological gambling prevalence estimates: facts or artefacts? *International Journal of Methods in Psychiatric Research*, 20(4), p.e83–e99.

Schüll, N.D., 2012. *Addiction by Design: Machine Gambling in Las Vegas*, Princeton, NJ: Princeton University Press.

Shaffer, H.J., 2005. From disabling to enabling the public interest: Natural transitions from gambling exposure to adaptation and self-regulation. *Addiction*, 100(9), p.1227–1230.

Shaffer, H.J., Hall, M.N. & Bilt, J.V., 1999. Estimating the prevalence of disordered gambling behavior in the United States and Canada: A research synthesis. *American Journal of Public Health*, 89(9), p.1369–1375.

Skog, O.-J., 1973. Less alcohol-fewer alcoholics. *The Drinking and Drug Practices Surveyor* 7, 7–14.

Skog, O.-J., 2006. The Historical Roots of Ledermann's Theory of the Distribution of Alcohol Consumption. *Contemporary Drug Problems* 33, p.143–174.

Smith, D.M., Clarke, G.P. & Harland, K., 2009. Improving the synthetic data generation process in spatial microsimulation models. *Environment and Planning A*, 41(5), p.1251–1268.

Smith, D.M., Pearce, J.R. & Harland, K., 2011. Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand. *Health & Place*, 17(2), p.618–624.

State Government of Victoria, 2014. Victorian Commission for Gambling Regulation: Expenditure by Venue. Available at: http://www.vcgr.vic.gov.au/CA256F800017E8D4/VCGLR/D6FBF3C7D7FC7E86CA257B3200786F30?OpenDocument archived at http://www.webcitation.org/6OTQismOM [Accessed March 31, 2014].

Stinchfield, R. & Winters, K.C., 2001. Outcome of Minnesota's gambling treatment programs. *Journal of Gambling Studies*, 17(3), p.217–245.

Stith, S.M. et al., 2004. Intimate partner physical abuse perpetration and victimization risk factors: A meta-analytic review. *Aggression and Violent Behavior*, 10(1), p.65–98.

Stockwell, T. & Gruenewald, P.J., 2003. Controls on the Physical Availability of Alcohol. In T. Stockwell & N. Heather, eds. *The essential handbook of treatment and prevention of alcohol problems*. New York: Wiley, J. Wiley, pp. 213–233.

Stone, C.A. et al., 2014. Effects of different screening and scoring thresholds on PGSI gambling risk segments. *International Journal of Mental Health and Addiction*, 13(1), p.82–102.

Storer, J., Abbott, M. & Stubbs, J., 2009. Access or adaptation? A meta-analysis of surveys of problem gambling prevalence in Australia and New Zealand with respect to concentration of electronic gaming machines. *International Gambling Studies*, 9(3), p.225–244.

St-Pierre, R.A. et al., 2014. How availability and accessibility of gambling venues influence problem gambling: A review of the literature. *Gaming Law Review and Economics*, 18(2), p.150–172.

Suomi, A. et al., 2013. Problem gambling and family violence: family member reports of prevalence, family impacts and family coping. *Asian Journal of Gambling Issues and Public Health*, 3(1), p.1–15.

Svetieva, E. & Walker, M., 2008. Inconsistency between concept and measurement: The Canadian Problem Gambling Index (CPGI). *Journal of Gambling Issues*, 22(2008), p.157–173.

Tanton, R. & Edwards, K.L., 2012. Limits of Static Spatial Microsimulation Models. In R. Tanton & K. Edwards, eds. *Spatial Microsimulation: A Reference Guide for Users*. Understanding Population Trends and Processes. Springer Netherlands, pp. 161–168.

The Economist, 2017. *The World's Biggest Gamblers*, London. Available at: http://www.economist.com/blogs/graphicdetail/2017/02/daily-chart-4 [Accessed April 1, 2017].

The World Bank, 2015. World Development Indicators. Available at: http://data.worldbank.org/data-catalog/world-development-indicators [Accessed April 29, 2015].

Thompson, S.G. & Higgins, J.P.T., 2002. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21(11), p.1559–1573.

Thorsen, I. & Gitlesen, J.P., 1998. Empirical evaluation of alternative model specifications to predict commuting flows. *Journal of Regional Science*, 38(2), p.273–292.

Tobler, W.R., 1990. Frame independent spatial analysis. In M. Goodchild & S. Gopal, eds. *Accuracy of spatial databases*. London: Taylor & Francis, pp. 115–122.

Turner, N. & Horbay, R., 2004. How do slot machines and other electronic gambling machines actually work? *Journal of Gambling Issues*, 11(2004), p.1–12.

Ugarte, M.D. et al., 2012. Gender-specific spatio-temporal patterns of colorectal cancer incidence in Navarre, Spain (1990–2005). *Cancer Epidemiology*, 36(3), p.254–262.

Vasiliadis, S.D. et al., 2013. Physical accessibility of gaming opportunity and its relationship to gaming involvement and problem gambling: A systematic review. *Journal of Gambling Issues*, 28(2013), p.1–46.

Victoria Police, 2014. *Code of Practice for the Investigation of Family Violence: Edition 3 V2 2014*, Melbourne. Available at: https://www.police.vic.gov.au/retrievemedia.asp?Media_ID=464 archived at http://www.webcitation.org/6XZyUwgbT [Accessed April 6, 2015].

Vine, M.M., Elliott, S.J. & Keller-Olaman, S., 2010. To disrupt and displace: placing domestic violence on the public health agenda. *Critical Public Health*, 20(3), p.339–355.

Walker, D.M., 2010. Casinos and crime in the USA. In B. L. Benson & P. R. Zimmerman, eds. *Handbook on the economics of crime*. Northampton, MA: Edward Elgar, pp. 488–517.

Walmsley, D.J. & Jenkins, J.M., 1992. Cognitive distance: A neglected issue in travel behavior. *Journal of Travel Research*, 31(1), p.24–29.

Wardle, H. et al., 2011. *British Gambling Prevalence Survey 2010*, London: National Centre for Social Research. Available at: http://www.gamblingcommission.gov.uk/pdf/british%20gambling%20prevalence%20survey%202010.pdf archived at http://www.webcitation.org/6b1OXuFhf [Accessed August 24, 2015].

Wardle, H. et al., 2016. *Exploring Area-Based Vulnerability to Gambling-Related Harm: Developing the Gambling-Related Harm Risk Index*, London: City of Westminster. Available at: http://transact.westminster.gov.uk/docstores/publications_store/licensing/final_phase2_exploring_area_based_vulnerability_to_gambling_related_harm.pdf archived at http://www.webcitation.org/6mMoxbRE8 [Accessed 1 January 2018].

Wardle, H. et al., 2014. 'Risky places?': Mapping gambling machine density and socio-economic deprivation. *Journal of Gambling Studies*, 30(1), p.201–212.

Weinstock, J., Ledgerwood, D.M. & Petry, N.M., 2007. Association between posttreatment gambling behavior and harm in pathological gamblers. *Psychology of Addictive Behaviors*, 21(2), p.185–193.

Welsh, M. et al., 2014. The 'problem gambler' and socio-spatial vulnerability. In F. Gobet & M. Schiller, eds. *Problem Gambling: Cognition, Prevention and Treatment*. London: Palgrave MacMillan, pp. 156–187. Available at: [Accessed February 5, 2015].

Welte, J.W. et al., 2002. Gambling participation in the U.S: Results from a national survey. *Journal of Gambling Studies*, 18(4), p.313–337.

Welte, J.W. et al., 2004. Risk factors for pathological gambling. *Addictive Behaviors*, 29(2), p.323–335.

Welte, J.W. et al., 2004. The relationship of ecological and geographic factors to gambling behavior and pathology. *Journal of Gambling Studies*, 20(4), p.405–423.

Wheeler, B.W., Rigby, J.E. & Huriwai, T., 2006. Pokies and poverty: Problem gambling risk factor geography in New Zealand. *Health & Place*, 12(1), p.86–96.

Wheeler, S.A. et al., 2008. The influence of gaming expenditure on crime rates in South Australia: A local area empirical investigation. *Journal of Gambling Studies*, 24(1), p.1–12.

Wheeler, S.A., Round, D.K. & Wilson, J.K., 2010. The relationship between crime and electronic gaming expenditure: Evidence from Victoria, Australia. *Journal of Quantitative Criminology*, 27(3), p.315–338.

Whitworth, A. et al., 2017. Estimating uncertainty in spatial microsimulation approaches to small area estimation: A new approach to solving an old problem. *Computers, Environment and Urban Systems*, 63(2017), p.50–57.

Wilkinson, J., 1996. *NSW and Gambling Revenue*, Sydney: NSW Parliamentary Library Research Service. Available at: http://www.parliament.nsw.gov.au/prod/parlment/publications.nsf/0/de920d27acd8d788ca256ecf00096910/$FILE/16-96.pdf archived at http://www.webcitation.org/6q2PfcoBB [Accessed April 1, 2017].

Williams, R.J. et al., 2015. *Quinte longitudinal study of gambling and problem gambling*, Ontario Problem Gambling Research Centre. Available at: http://hdl.handle.net/10133/3641 archived at http://www.webcitation.org/6b1SWDd0o [Accessed August 24, 2015].

Williams, R.J. & Volberg, R.A., 2009. Impact of survey description, administration format, and exclusionary criteria on population prevalence rates of problem gambling. *International Gambling Studies*, 9(2), p.101–117.

Williams, R., Volberg, R., 2010. *Best practices in the population assessment of problem gambling*. Ontario Problem Gambling Research Centre, Guelph. Available at: http://hdl.handle.net/10133/1259 [Accessed July 1, 2016].

Williams, R.J. & Volberg, R.A., 2013. The classification accuracy of four problem gambling assessment instruments in population research. *International Gambling Studies*, 14(1), p.15–28.

Williams, R.J., Volberg, R.A. & Stevens, R.M.G., 2012. *The Population Prevalence of Problem Gambling: Methodological Influences, Standardized Rates, Jurisdictional Differences, and Worldwide Trends*, Ontario Problem Gambling Research Centre. Available at: http://www.uleth.ca/dspace/handle/10133/3068 archived at http://www.webcitation.org/6OK7CeYsW [Accessed March 21, 2014].

Williams, R.J. & Wood, R.T., 2004. The proportion of gaming revenue derived from problem gamblers: Examining the issues in a Canadian context. *Analyses of Social Issues & Public Policy*, 4(1), p.33–45.

Wood, R.T. & Williams, R.J., 2007. 'How much money do you spend on gambling?' The comparative validity of question wordings used to assess gambling expenditure. *International Journal of Social Research Methodology*, 10(1), p.63–77.

Wood, R.T.A. & Griffiths, M.D., 2007. A qualitative investigation of problem gambling as an escape-based coping strategy. *Psychology and Psychotherapy: Theory, Research and Practice*, 80(1), p.107–125.

Woolley, R. & Livingstone, C., 2010. Into the zone: Innovating in the Australian poker machine industry. In S. Kingma, ed. *Global Gambling—Cultural Perspectives on Gambling Organizations*. London; New York: Routledge, pp. 38–63.

Yakovenko, I., Hodgins, D.C., 2018. A scoping review of co-morbidity in individuals with disordered gambling. *International Gambling Studies,* 18(1), p.143–172.

Young, M. et al., 2006. *Northern Territory Gambling Prevalence Survey*, Darwin: School for Social and Policy Research, Charles Darwin University. Available at: https://justice.nt.gov.au/licensing-nt-and-nt-worksafe/licensing-publications-and-strategies/gambling-research/northern-territory-gambling-prevalence-survey-2005 archived at http://www.webcitation.org/6q2PwSIVC [Accessed April 28, 2017].

Young, M., 2013. Statistics, scapegoats and social control: A critique of pathological gambling prevalence research. *Addiction Research & Theory*, 21(1), p.1–11.

Young, M., Lamb, D. & Doran, B., 2009. Mountains and molehills: A spatiotemporal analysis of poker machine expenditure patterns in the Northern Territory. *Australian Geographer*, 40(3), p.249–269.

Young, M., Markham, F., 2015. Beyond disclosure: gambling research, political economy, and incremental reform. *International Gambling Studies* 15, 6–9.

Young, M., Markham, F. & Doran, B., 2012a. Placing bets: Gambling venues and the distribution of harm. *Australian Geographer*, 43(4), p.425–444.

Young, M., Markham, F. & Doran, B., 2012b. Too close to home? The relationships between residential distance to venue and gambling outcomes. *International Gambling Studies*, 12(2), p.257–273.

Young, M., Stevens, M. & Morris, M., 2008. Problem gambling within the non-Indigenous population of the Northern Territory of Australia: A multivariate analysis of risk factors. *International Gambling Studies*, 8(1), p.77–93.

Ziolkowski, S., 2016. *World Count of Gaming Machines 2015*, Sydney: Gaming Technologies Association. Available at: http://gamingta.com/wp-content/uploads/2016/10/World_Count_of_Gaming_Machines_2015.pdf archived at http://www.webcitation.org/6pzANhiht [Accessed February 24, 2017].

Appendix A: Declaration of authorship

Chapters 2 – 7 in this thesis consist of co-authored research papers. This appendix lists the status of the papers at the time of submission and the attribution of work among the authors. In each case, Markham was the lead author.

Chapter 2

Publication

Markham, F., Doran, B. & Young, M., 2014. Estimating gambling venue catchments for impact assessment using a calibrated gravity model. *International Journal of Geographical Information Science*, 28(2), p.326–342.

Status

Published.

Attribution statement

All authors jointly conceived of this study. MY and BD designed the questionnaire. MY supervised data collection. FM designed and undertook data analysis and interpreted the results. FM drafted the manuscript. All authors contributed to critically revising the manuscript. All authors approved the manuscript for publication.

Chapter 3

Publication

Markham, F., Young, M. & Doran, B., 2016. The relationship between player losses and gambling-related harm: evidence from nationally representative cross-sectional surveys in four countries. *Addiction*, 111(2), p.320–330.

Status

Published.

Attribution statement

FM conceived of this study. FM led the design of the study, with contributions from MY and BD. FM designed and undertook data analysis. FM led the interpretation of the results, with contributions from MY and BD. FM drafted the manuscript. All authors contributed to critically revising the manuscript. All authors approved the manuscript for publication.

Chapter 4

Publication

Markham, F., Young, M. & Doran, B., 2014. Gambling expenditure predicts harm: evidence from a venue-level study. *Addiction*, 109(9), p.1509–1516.

Status

Published.

Attribution statement

FM conceived of this study. MY and BD designed the questionnaire. MY supervised data collection. FM led the design of the study, with contributions from MY and BD. FM designed and undertook data analysis. FM led the interpretation of the results, with contributions from MY and BD. FM drafted the manuscript. All authors contributed to critically revising the manuscript. All authors approved the manuscript for publication.

Chapter 5

Publication

Markham, F., Young, M., Doran, B. & Sugden, M. [Currently unpublished.] A meta-regression analysis of 41 Australian problem gambling prevalence estimates and their relationship to total spending on electronic gaming machines. Accepted for publication in *BMC Public Health*.

Status

Accepted for publication by editor Noriko Cable in *BMC Public Health* on April 20, 2017, but not yet published at time of thesis submission.

Attribution statement

FM conceived of this study. All authors contributed to the design the data collection instrument. FM and MS conducted systematic literature searches and extracted data from eligible studies. MY contributed data extraction where FM and MS disagreed about coding. FM led the design of the study, with contributions from MY and BD. FM designed and undertook data analysis. FM led the interpretation of the results, with contributions from MY and BD. FM drafted the manuscript. All authors contributed to critically revising the manuscript. All authors approved the manuscript for publication.

Chapter 6

Publication

Markham, F., Young, M. & Doran, B. [Currently unpublished.] Improving spatial microsimulation estimates of health outcomes by including geographic indicators of health behaviour: The example of problem gambling. Accepted for publication in *Health and Place.*

Status

Accepted for publication by editor Jamie Pearce in *Health and Place* on April 25, 2017, but not yet published at time of thesis submission. Doi when published: 10.1016/j.healthplace.2017.04.008

Attribution statement

All authors jointly conceived of this study. MY and BD designed the questionnaire. MY supervised data collection. FM designed and undertook data analysis and interpreted the results. FM drafted the manuscript. All authors contributed to critically revising the manuscript. All authors approved the manuscript for publication.

Chapter 7

Publication

Markham, F., Doran, B. & Young, M., 2016. The relationship between electronic gaming machine accessibility and police-recorded domestic violence: A spatio-temporal analysis of 654 postcodes in Victoria, Australia, 2005–2014. *Social Science & Medicine*, 162, p.106–114.

Status

Published

Attribution statement

FM conceived of this study. FM designed and undertook data analysis. All authors contributed to interpreting the results. FM drafted the manuscript. All authors contributed to critically revising the manuscript. All authors approved the manuscript for publication.

Author signatures

I agree that the declarations above provide an accurate description of my contribution to the manuscripts presented in this thesis.

Francis Markham 25 April 2017

Bruce Doran 25 April 2017

Martin Young 25 April 2017

Mark Sugden 25 April 2017

Appendix B: Supplementary tables for Chapter 2

In order to estimate the effect of distance from their residence to a gamer's preferred gaming venue accurate geospatial measures of the location of residences is required. The most feasible, cost-effective way of producing a sample of geocoded responses was to conduct a mail-out survey using an existing geocoded address database. The Geocoded National Address File (G-NAF) produced by the Public Sector Mapping Agencies (PSMA) Australia provided the means to conduct such a survey. G-NAF contains the latitude, longitude, suburb, street and house number for addresses in Australia, derived from the combined databases of mapping agencies of federal and state governments, Australia Post and the Australian Electoral Commission (PSMA Australia 2010). A version of the February 2010 G-NAF was licensed which included zoning category (e.g. Residential, Industrial, Education, Agricultural, etc.) from the Australian Bureau of Statistics' 2006 mesh blocks, zoning categories that were themselves derived from state government sources (Australia Bureau of Statistics 2005, p. 5). All questionnaires were delivered with a reply-paid envelope and a web-address where the survey could alternatively be completed. Only 159 respondents or 2.2% elected to submit the survey electronically. Each questionnaire contained a unique identifier that was used to geocode the residence to which it was delivered.

Surveys were sent to all G-NAF addresses in the Northern Territory to which Australia Post will deliver unsolicited mail to street addresses and which were in a 'residential' planning zone, as defined by the ABS 2006 mesh blocks. In addition, residential addresses in selected 'industrial' zones, most notably Darwin's Central Business District (CBD), were included on the basis of local knowledge. This resulted in 46,288 surveys being delivered to households in Darwin, Katherine and Alice Springs between April and August 2010 (see Table 1), with a response rate of 14.3%.

In addition, 3,465 addresses outside of Australia Post's delivery zone were selected for hand delivery of surveys on the peri-urban fringes of Alice Springs and Darwin in order to increase the spatial coverage of the survey (see Table 1). In Alice Springs' peri-urban fringe, questionaires were delivery to only 20% of the selected rural addresses, as many had no residence or were unidentifiable. Hand deliveries in Alice Springs were to addresses within 15km of the CBD, and took place in July 2010.

In the Darwin peri-urban fringe, 2,000 out of a potential 7,000 addresses between 20km and 40km from the CBD were selected for hand delivery of surveys in September 2010. The peri-urban fringe was divided in bands based on distance to the CBD with a resolution of 5km. In each band 500 addresses were selected for delivery, in several contiguous blocks (see Figure 1). In total, 2,300 surveys were handed delivered with a response rate of 19.7%.

Delivery mode	*Location*	*Date conducted*	*Questionnaires mailed*	*Questionnaires successfully delivered*	*Questionnaires returned*	*Response rate*
Postal delivery	Alice Springs	23/04/2010	10,049	8,571	1,313	15.3%
	Katherine	04/08/2010	3,058	1,894	205	10.8%
	Darwin	31/04/2010	42,298	35,798	5,078	14.2%
	Subtotal		**55,405**	**46,263**	**6,596**	**14.3%**
Hand delivery	Darwin	14-24/09/2010	NA	2,000	374	18.7%
	Alice Springs	24-5/07/2010	NA	300	79	26.3%
	Subtotal			**2,300**	**453**	**19.7%**
Total			**55,405**	**48,563**	**7,049**	**14.5%**

Table B.1: Survey response rates

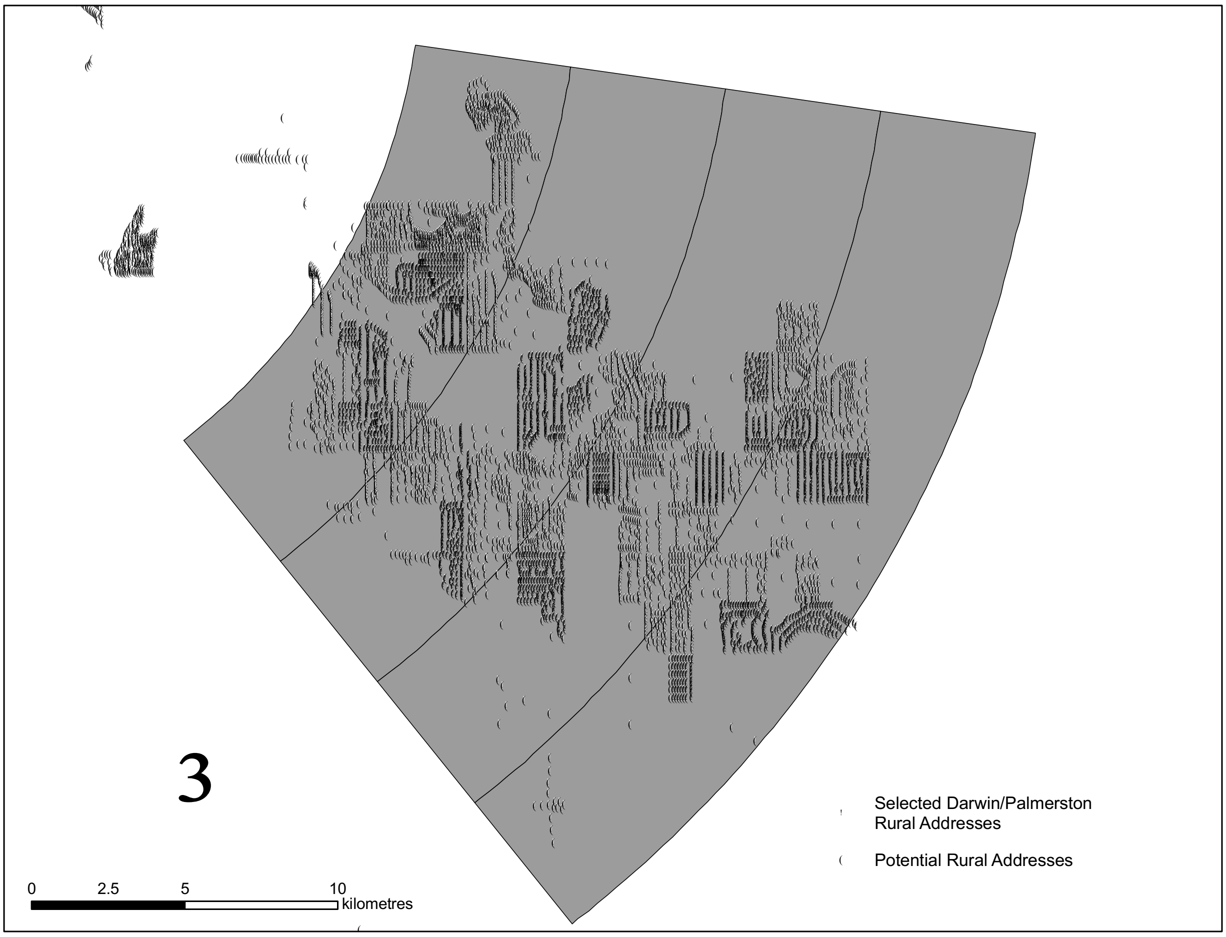

Figure B.1: Darwin peri-urban sample design

Appendix C: Supplementary figures and tables for Chapter 3

C.1 Supplementary results

C.1.1 Smaller x-axis width

It is possible that a curve that appears to be linear or r-shaped may have a J-shaped inflection at relatively moderate levels of gambling losses that is policy relevant. Consequently, figures in the main text have been replicated here as Figures C.1 and C.2, but with their x-axes truncated at $250 USD per month. In other words, Figures C.1 and C.2 show the estimated curve shape for the first $250 USD per month of losses, rather than the full data range.

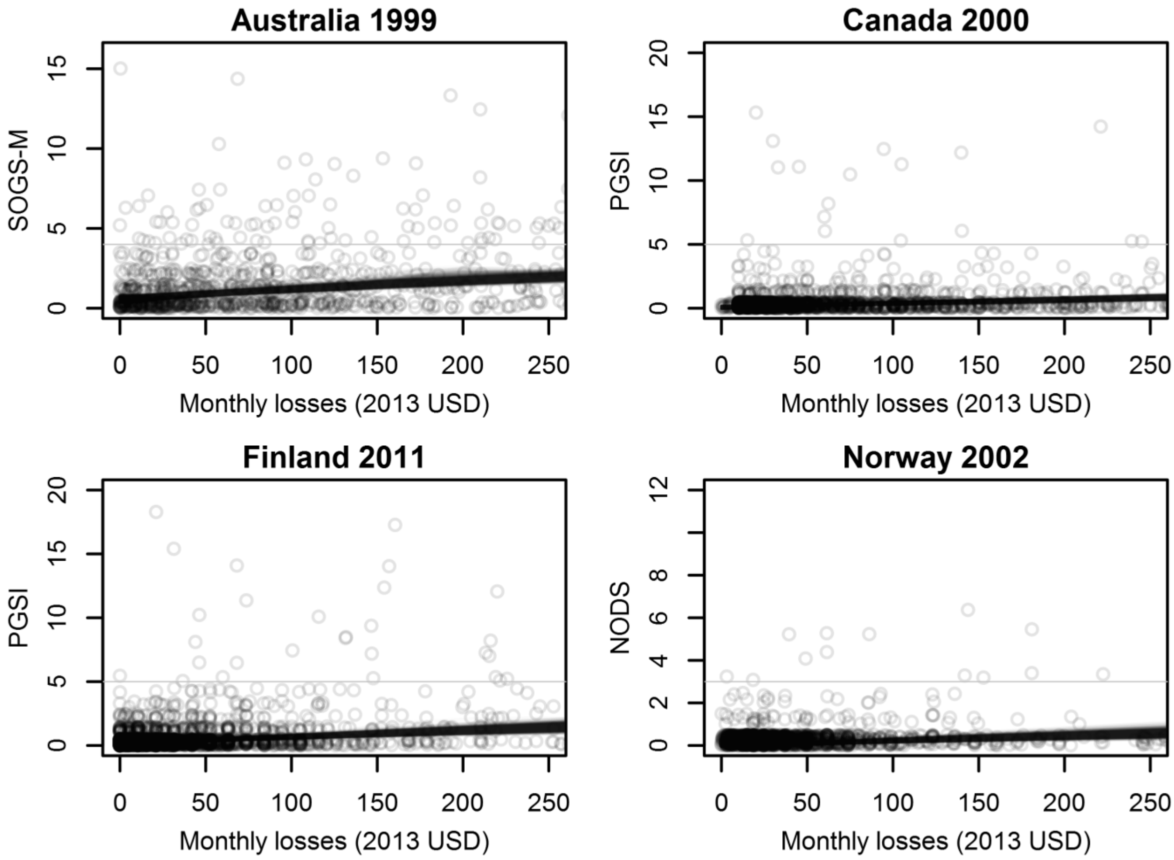

Figure C.1: Bootstrapped risk curves for total gambling losses versus problem gambling risk, with plot constrained to the range $0 - $250 USD per month. Horizontal lines represent the standardised problem gambling thresholds calculated by Williams & Volberg (2013): SOGS-M = 4; PGSI = 5; NODS = 3. Losses are standardised to 2013 US dollars spent in previous 30 days. Each point represents a single respondent, jittered for display. Each line represents a single non-parametric bootstrapped loess fit, with span selected by AICc. Median spans [and 95% CIs] were: 1.0 [0.4, 5.0], 1.0 [0.8, 5.0], 1.0 [0.6, 5.0] and 1.4 [0.6, 5.0].

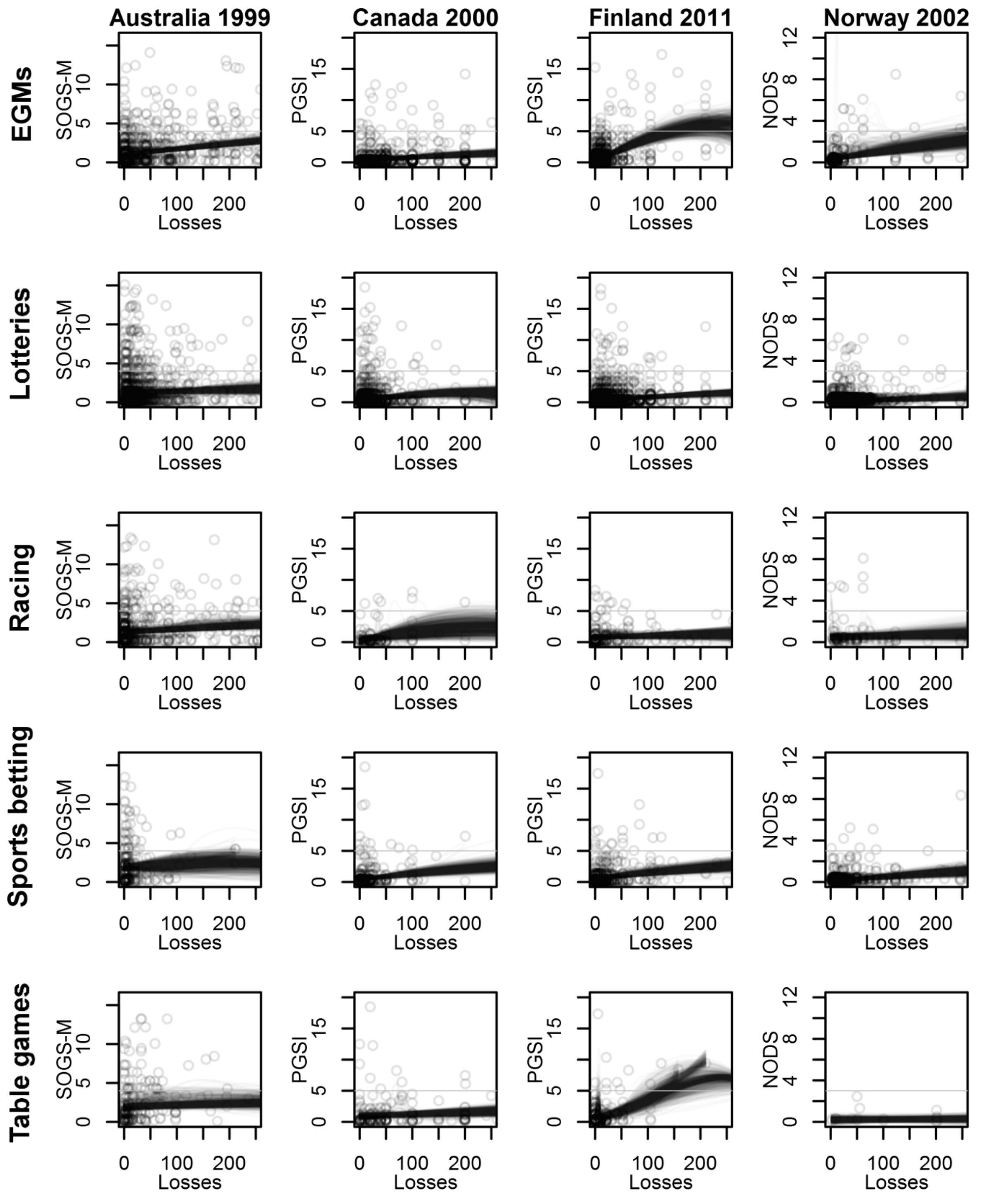

Figure C.2 Bootstrapped risk curves for gambling losses versus problem gambling risk, with plot constrained to the range $0 - $250 USD per month. Horizontal lines represent the standardised problem gambling thresholds calculated by Williams & Volberg (2013): SOGS-M = 4; PGSI = 5; NODS = 3. Losses are standardised to 2013 US dollars spent in previous 30 days. Each point represents a single respondent, jittered for display only.

C.1.2 Non-bootstrapped analysis

Figure 3.2 and Figure 3.3 in the main text used a loess smoother to examine the relationship between losses and harm using non-parametric bootstrap to minimise the impact of endpoints on curve shape. Figures C.3 and C.4 replicate the analyses presented in Figure 3.2 and Figure 3.3 in the main text, but without using bootstrap methods. No evidence of J-shaped curves is readily apparent.

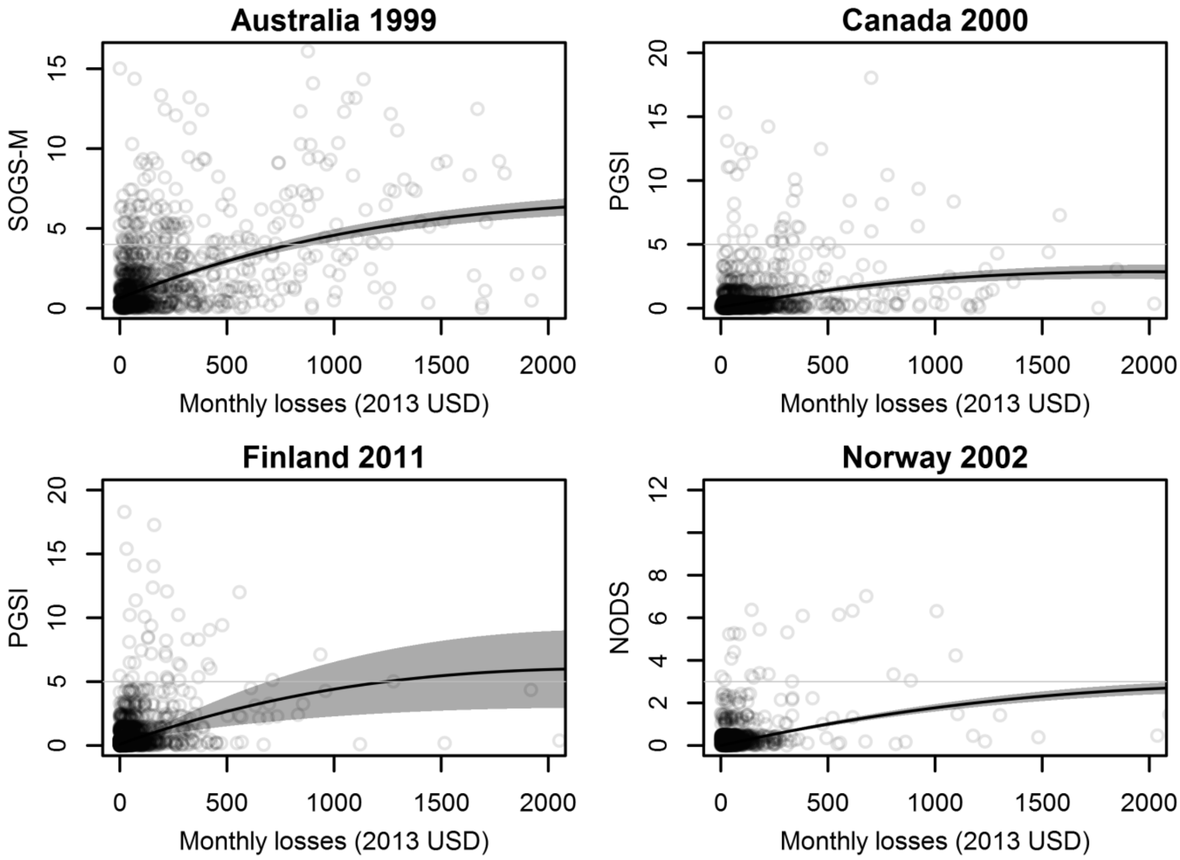

Figure C.3: Risk curves for total gambling losses versus problem gambling risk. The shaded region represents Wald-style 95% confidence intervals. Horizontal lines represent the standardised problem gambling thresholds calculated by Williams & Volberg (2013): SOGS-M = 4; PGSI = 5; NODS = 3. Losses are standardised to 2013 US dollars spent in previous 30 days. Each point represents a single respondent, jittered for display.

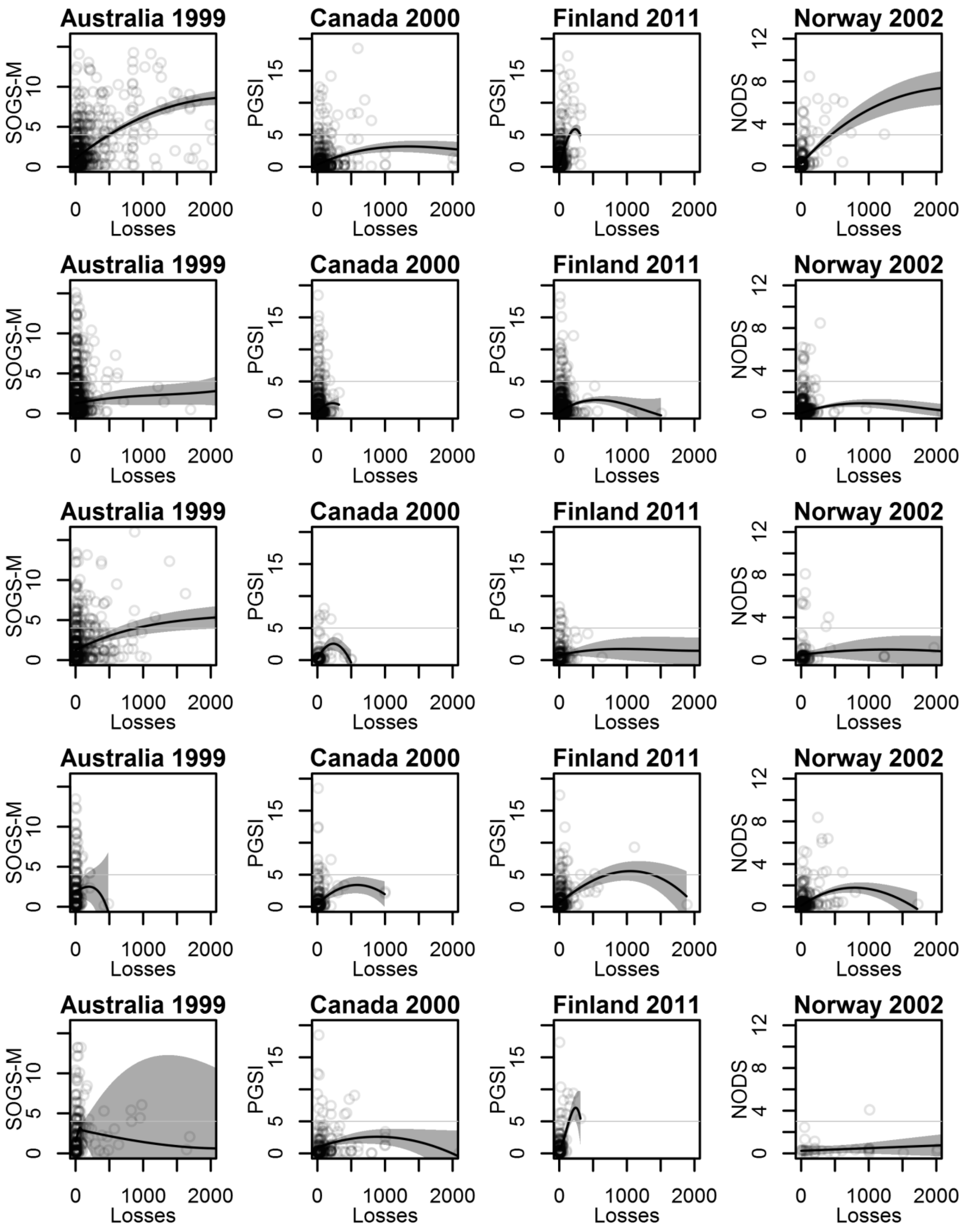

Figure C.4: Risk curves for gambling losses versus problem gambling risk for five gambling activities. The shaded region represents Wald-style 95% confidence intervals. Horizontal lines represent the standardised problem gambling thresholds calculated by Williams & Volberg (2013): SOGS-M = 4; PGSI = 5; NODS = 3. Losses are standardised to 2013 US dollars spent in previous 30 days. Each point represents a single respondent, jittered for display only.

C.2 Population weighted analysis

Table 3.2 in the main text has been reanalysed using population weights. The results are consistent with those estimated without population weights in the main text.

Table C.1 Multiple linear regression estimates of player loss – problem gambling risk curves by gambling activity, estimated using population weights.

		Australia 1999 [95% confidence interval]	Canada 2000 [95% confidence interval]	Finland 2011 [95% confidence interval]	Norway 2002 [95% confidence interval]
Total	$10^3 \times \beta$ *losses*	**4.6 [3.2, 6.6]**	**1.7 [0.9, 3.5]**	**3.7 [2.3, 8.1]**	**1.7 [0.7, 3.3]**
	$10^7 \times \beta$ *losses* ^{2}	**-8.5 [-17.5, -4.4]**	**-3.1 [-14.7, -1.4]**	**-4.1 [-35.8, -2.0]**	-2.9 [-13.9, 1.1]
	losses R^2	0.21 [0.13, 0.31]	0.04 [0.01, 0.12]	0.11 [0.07, 0.19]	0.13 [0.07, 0.25]
	n	896	1259	3004	1875
EGMs	$10^3 \times \beta$ *losses*	**6.9 [5.2, 12.1]**	**3.1 [1.3, 8.0]**	**39.6 [23.5, 56.0]**	**5.1 [2.7, 20.2]**
	$10^7 \times \beta$ *losses* ^{2}	**-14.9 [-53.8, -10.0]**	**-7.4 [-37.9, -2.4]**	-672.0 [-1365.2, 89.6]	**-9.1 [-315.0, -2.1]**
	losses R^2	0.25 [0.16, 0.37]	0.04 [0.00, 0.18]	0.19 [0.04, 0.32]	0.21 [0.15, 0.49]
	n	619	462	1156	180
Lotteries	$10^3 \times \beta$ *losses*	3.9 [-3.3, 9.3]	**17.2 [7.5, 27.7]**	**6.8 [4.5, 11.6]**	3.1 [-1.2, 5.4]
	$10^7 \times \beta$ *losses* ^{2}	-33.7 [-146.0, 150.0]	**-745.0 [-1294.8, -224.0]**	**-48.1 [-255.0, -30.2]**	-31.8 [-60.2, 279.0]
	losses R^2	0.01 [0.00, 0.06]	0.03 [0.01, 0.06]	0.03 [0.01, 0.08]	0.03 [0.00, 0.09]
	n	722	1073	2700	1943
Racing	$10^3 \times \beta$ *losses*	**3.0 [0.6, 6.9]**	2.9 [-11.9, 26.9]	-1.5 [-4.4, 7.1]	0.8 [-1.0, 7.6]
	$10^7 \times \beta$ *losses* ^{2}	-2.7 [-24.2, 8.0]	-118.0 [-1151.8, 412.0]	7.9 [-79.6, 17.7]	-3.3 [-59.4, 6.5]
	losses R^2	0.07 [0.01, 0.19]	0.06 [0.02, 0.75]	0.12 [0.00, 0.42]	0.14 [0.03, 0.64]
	n	453	68	215	101
Sports	$10^3 \times \beta$ *losses*	0.6 [-26.0, 19.5]	3.8 [-7.4, 9.9]	**10.0 [3.3, 15.5]**	**4.5 [1.9, 10.2]**

betting	$10^7 \times \beta$ *losses* ^{2}	-92.9 [-565.0, 1747.8]	-17.9 [-79.6, 392.0]	-48.2 [-213.0, 38.5]	**-28.1 [-133.0, -10.9]**
	losses R^2	0.10 [0.03, 0.34]	0.07 [0.03, 0.24]	0.06 [0.00, 0.25]	0.14 [0.07, 0.29]
	n	175	222	435	402
Table games	$10^3 \times \beta$ *losses*	3.9 [-3.3, 11.4]	1.2 [-4.1, 3.8]	34.6 [-13.0, 55.8]	0.2 [-0.8, 3.4]
	$10^7 \times \beta$ *losses* ^{2}	-29.1 [-151.0, 61.2]	-6.8 [-34.4, 58.8]	-319.0 [-1161.2, 4153.5]	0.0 [-16.4, 3.5]
	losses R^2	0.00 [0.00, 0.19]	0.07 [0.01, 0.33]	0.13 [0.00, 0.49]	0.11 [0.00, 0.91]
	n	169	126	172	31

Notes: Player loss β coefficients estimated from multiple linear regression. Square brackets report 95% confidence intervals, estimated via the percentile method from an ordinary, non-parametric bootstrap with 5,000 replications. Estimated coefficients are not reported for socio-demographic predictor variables for reasons of brevity. Non-reported predictor variables include: age; age^2; sex; education level; marital status; employment status; household income; and household $income^2$. *Losses* R^2 reports the variance explained by the player loss terms in the regression, after adjusting for other covariates. *Losses* R^2 was calculated by subtracting the adjusted R^2 of the full multiple linear regression from that of a multiple linear regression specified identically except with the player loss terms dropped. † indicates that due to the small number of observations, the linear regression was specified without socio-demographic predictor variables.

C.2: Risk curves and problem gamblers' loss share

Introduction

It is well established that problem gamblers contribute a disproportionate amount of gambling losses, especially on electronic gaming machines (EGMs). For example, the Australian Productivity Commission synthesised 21 estimates of the loss share of problem gamblers on EGMs in Australia, and found that, on average, problem gamblers account for 41 per cent of losses (Productivity Commission 2010, p.C.11). Given that the same report estimated the prevalence of problem gambling among Australian adults to be just 0.85% (Productivity Commission 2010, p.5.23), it is clear that problem gamblers lose much more money than other gamblers.

It is intuitive that such a relationship is compatible with a J-shaped relationship between gambling losses and gambling-related harm. However, while a J-shaped relationship is a sufficient for problem gamblers to account for a disproportionate share of losses, it is not a necessary condition.

This appendix uses a simple simulation to demonstrate that a linear or r-shaped relationship between losses and harm still may result in problem gamblers losing a disproportionate amount of money.

Methods

A population of 100,000 gamblers was generated using parameters from the Australian Productivity Commission's 2010 report. (Productivity Commission 2010) Each simulated gambler had two features of interest: (a) A score on the Problem Gambling Severity Index (PGSI), and (b) the amount lost on EGMs in the last year. The simulated population was constructed of four subgroups (see Table C.2). Simple assumptions were used to assign PGSI scores and losses for transparency and consistency with the Productivity Commission data. Within each of the four PGSI bands, we applied the perhaps overly stringent, conservative assumption that the relationship between expenditure and harm is random.

A simplified loess and regression analysis was carried out on this simplified data set. No bootstrapping was applied. A loess span of 1.01 was selected by minimising AICc

(Hurvich et al. 1998). The R code used to produce this simulation is presented in Listing C.1.

Results

An r-shaped relationship was found in this simulation study, both in the loess non-parametric curve fitting (Figure C.5) and in the quadratic bivariate regression (Table C.3). As Figure C.5 shows, this simulation is highly simplified and its parameters are only coarse approximations of these relationships in the general population. In particular, the distributional assumptions used to estimate PGSI scores and losses within each of the four bands are unrealistic. Nevertheless, even this simple simulation strongly demonstrates that a linear or r-shaped harm-loss relationship is compatible with a high problem-gambler expenditure share.

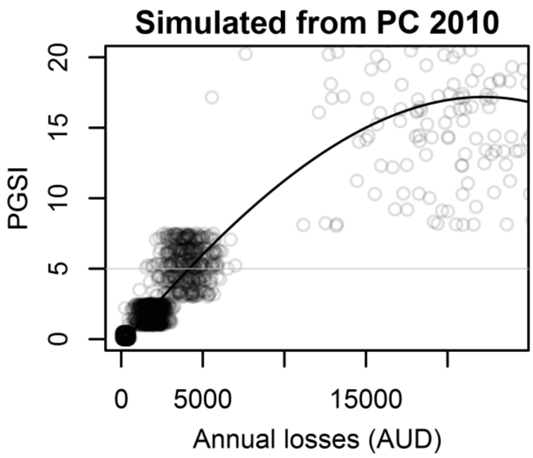

Figure C.5: The gambling-related harm – gambling expenditure curve in a simulated population of gamblers where those with PGSI ≥ 8 account for 42% of gambling losses.

Table C.2: Characteristics of the simulated population of gamblers.

	Non-problem gamblers	Low risk gamblers	Moderate risk gamblers	Problem gamblers
n (%) [a]	8,540 (85.4%)	810 (8.1%)	440 (4.4%)	200 (2.0%)
Distribution of PGSI scores	Fixed at 0	Uniform distribution from 1 to 2	Uniform distribution from 3 to 7	Uniform distribution from 8 to 27
Distribution of losses [b]	Normal distribution with mean = \$274 sd = \$68.50	Normal distribution with mean = \$1879 sd = \$469.75	Normal distribution with mean = \$4059 sd = \$1014.75	Normal distribution with mean = \$20625 sd = \$5156.25
Share of losses [c]	24.1%	15.6%	18.1%	42.1%

Notes:
[a] Percentages sourced from the mean of estimates of the proportion of EGM gamblers in each PGSI category in seven Australian state prevalence studies (NSW 2006, Tas 2007, Qld 2006-7, Qld 2008-8, SA 2005, Vic 2003, Vic 2008) in the Productivity Commission report's *Appendix B*.
[b] Mean losses for each PGSI category, averaged across 21 estimates produced by the Productivity Commission from seven Australian state prevalence studies, presented in the reports *Appendix B*.
[c] Estimated in the simulation but consistent with proportions in the Productivity Commission's 2010 report.

Table C.3: Bivariate quadratic linear regression estimates of player loss – problem gambling risk curves in simulation.

	Estimate	95% confidence interval	p
Constant $\times 10^{-1}$	-4.5	[-4.8, -4.3]	< 0.0001
$10^3 \times \beta$ losses	1.5	[1.5, 1.5]	< 0.0001
$10^7 \times \beta$ losses ^{2}	-3.1	[-3.1, -3.0]	< 0.0001

Notes: n = 10,000, R^2 = 0.87. Standard errors are small due to the simulated nature of the data.

Listing C.1: R code used to produce the simulation.

```
library(dplyr)

N.sims <- 10000

#### Generate simulation data ####
sims <- rbind(data.frame(PG.status=rep("Recreational",
round(N.sims*0.854)),
                         PGSI=rep(0, round(N.sims*0.854)),
                         Loss=rnorm(n = round(N.sims*0.854),
                                    mean = 274, sd=274/4)),
              data.frame(PG.status=rep("Low risk", round(N.sims*0.081)),
                         PGSI=round(runif(round(N.sims*0.081),
                                          min = 1.0, max = 2.0)),
                         Loss=rnorm(n = round(N.sims*0.081),
                                    mean = 1879, sd= 1879/4)),
              data.frame(PG.status=rep("Moderate risk",
round(N.sims*0.044)),
                         PGSI=round(runif(round(N.sims*0.044),
                                          min = 2.51, max = 7.49)),
                         Loss=rnorm(n = round(N.sims*0.044),
                                    mean = 4059, sd=4059/4)),
              data.frame(PG.status=rep("High risk", round(N.sims*0.020)),
                         PGSI=round(runif(round(N.sims*0.020),
                                          min = 7.51, max = 27.49)),
                         Loss=rnorm(n = round(N.sims*0.020),
                                    mean = 20625, sd=20625/4)))

# Calculate loss shares for population
group_by(sims, PG.status) %>%
  summarise(Total_loss=sum(Loss)) %>%
  mutate(Share=round((Total_loss/sum(Total_loss))*1000)/10) %>% print

#### Plot non-parametric smoother graph ####
pred.lines <- function(fit, xs, points=200, alpha=0.25){
  x.pts <- seq(from=min(xs), to=max(xs), length.out=points)
  lines(x=x.pts, y=predict(object = fit, newdata = x.pts),
        col=rgb(0, 0, 0, alpha))
}

# source('autoloess.R')
# Loess with automatic span determination
# lo <- autoloess(loess(PGSI ~ Loss, data=sims), span=c(0.1, 5.0))
# Result: Span = 1.012615
lo <- loess(PGSI ~ Loss, data=sims, span=1.012615)

# Jitter points for display only
sims <- mutate(sims, PGSIj=PGSI + runif(nrow(sims), min = 0, max = 0.5),
               LossJ=Loss + runif(nrow(sims), min = 0, max = Loss/100))

plot(x=sims$LossJ, y=sims$PGSIj, col=rgb(1, 1, 1, 0),
     xlab="Annual losses (AUD)", ylab="PGSI",
     main="Simulated from PC 2010",
     ylim=c(0, 20), xlim=c(0, 24000))
pred.lines(lo, sims$Loss, alpha=1)
points(x=sims$LossJ, y=sims$PGSIj, col=rgb(0, 0, 0, 1/8))
abline(h=5, lwd=0.4, col=rgb(0.75, 0.75, 0.75))

#### Quadratic bivariate regression ####
lm.fit <- lm(data=sims, PGSI ~ Loss + I(Loss^2))
print(summary(lm.fit))
print(cbind(coef(lm.fit), confint(lm.fit)))
```

C.3: Missing data

C.3.1 Introduction

Incomplete data may bias the estimation of curve shape if data are not missing completely at random (Little et al. 2012).

A substantial proportion of data was missing in each analysis, depending on the survey. For the Australian 1999 survey, a large proportion of the respondents failed to complete the household income question on the survey, resulting in around 30% of them being excluded from the full model. In Finland, Norway and Canada the typical number of respondents with missing data was closer to 13%, 14% and 40%, respectively.

C.3.2 Methods

Therefore, sensitivity analyses were conducted to determine if the non-response was likely to be biasing the results presented in the main text. For each regression analysis presented in the main text, we estimated four models:

1. The full model presented in the text, with responses with missing data removed listwise
2. The full model presented in the text, with missing data imputed by multiple imputation
3. A reduced model, with covariates only for age and sex, responses with missing data removed listwise
4. A minimal model, with no covariates and with responses with missing data removed listwise

Multiple imputation was carried out using chained equations with Gibbs sampling, with 10 imputed data sets for each analysis. All analyses were conducted using R (R Core Team 2015; van Buuren & Groothuis-Oudshoorn 2011).

For computational reasons, we did not apply the bootstrap to these sensitivity analyses. Consequently, the confidence intervals are overly tight. Interpretation should bear this in mind.

C.3.3 Results

Detailed results are presented in tables C.4 to C.26 (table games in Norway was excluded due to the small number of respondents who played them). To summarise, of the 23 relationships tested, only four changed from significant to non-significant or vice versa with multiple imputation.

While clearly changing from significance to non-significance is a problematic criteria (e.g. Gelman & Stern 2006), our interpretation of curve shape is based on this so it is a meaningful guide to how our interpretation might change on the basis of missing data.

The changed parameters are:

1. Multiple imputation (MI) found only a linear relationship for total losses in Australia, listwise exclusion (LE) found an r-shaped curve
2. MI found no relationship for lotteries in Australia, LE found a J-shaped curve
3. MI found only a linear relationship for lotteries in Canada, LE found an r-shaped curve
4. MI found an r-shaped curve for racing in Canada, LE found no relationship

Table C.4: Missing data sensitivity analysis for Australia, total losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.76	-1.41	2.92	0.03	-2.57	2.63	2.17	1.15	3.18	1.20	1.06	1.34
Losses x 10^3	2.43	1.89	2.97	3.23	2.67	3.80	2.51	2.06	2.95	2.53	2.08	2.98
Losses^2 x 10^7	-1.64	-3.62	0.34	-3.35	-5.26	-1.45	-1.79	-3.36	-0.23	-1.74	-3.32	-0.16
Age	0.00	-0.05	0.06	0.02	-0.05	0.08	-0.01	-0.06	0.03			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	-0.07	-0.34	0.21	-0.11	-0.42	0.21	-0.09	-0.35	0.17			
Employment - Other	-0.93	-2.59	0.73	-2.75	-5.05	-0.44						
Employment - Retired or pensioner	-0.18	-0.86	0.50	-0.22	-1.05	0.62						
Employment - Student	0.66	-0.31	1.63	1.38	0.11	2.65						
Employment - Unemployed	0.79	-0.14	1.73	0.39	-0.71	1.50						
Employment - Working fulltime or parttime	-0.03	-0.60	0.54	-0.20	-0.88	0.47						
Education - School only	0.94	-0.63	2.50	1.32	-0.66	3.30						
Education - Technical	1.13	-0.46	2.73	1.66	-0.34	3.66						
Education - University	0.93	-0.66	2.53	1.30	-0.71	3.30						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital_status - Separated or divorced	0.39	-0.07	0.86	0.28	-0.25	0.80						
Marital_status - Single	0.26	-0.14	0.65	0.36	-0.08	0.80						
Marital_status - Widowed	0.08	-0.49	0.64	0.21	-0.45	0.88						
n	1270			896			1225			1228		
Missing	0%			29%			4%			3%		
adj. R^2	0.19			0.23			0.19			0.17		

Table C.5: Missing data sensitivity analysis for Australia, EGM losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.64	-1.90	3.18	0.14	-3.13	3.42	2.43	1.22	3.64	1.21	1.03	1.39
Losses x 10^3	5.74	4.93	6.55	6.41	5.48	7.35	5.71	4.91	6.52	5.78	4.97	6.59
$Losses^2$ x 10^7	-11.72	-14.94	-8.50	-13.15	-16.57	-9.72	-11.73	-14.97	-8.50	-11.82	-15.09	-8.56
Age	-0.02	-0.08	0.05	-0.02	-0.10	0.06	-0.02	-0.08	0.03			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	-0.25	-0.58	0.08	-0.17	-0.55	0.21	-0.29	-0.61	0.02			
Employment - Other	-1.30	-3.42	0.82	-4.37	-7.08	-1.66						
Employment - Retired or pensioner	-0.33	-1.11	0.46	-0.63	-1.63	0.37						
Employment - Student	0.36	-0.74	1.45	0.79	-0.65	2.23						
Employment - Unemployed	0.20	-1.03	1.42	-0.96	-2.35	0.42						
Employment - Working fulltime or parttime	-0.12	-0.80	0.56	-0.42	-1.24	0.40						
Education - School only	1.73	-0.20	3.66	2.57	-0.04	5.18						
Education - Technical	1.83	-0.14	3.80	2.77	0.14	5.40						
Education - University	1.93	-0.02	3.88	2.92	0.29	5.54						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital_status - Separated or divorced	0.30	-0.28	0.88	0.27	-0.37	0.91						
Marital_status - Single	0.10	-0.36	0.57	0.26	-0.28	0.79						
Marital_status - Widowed	0.03	-0.63	0.69	-0.13	-0.93	0.67						
n	894			619			846			849		
Missing	0%			31%			5%			5%		
adj. R^2	0.27			0.30			0.27			0.25		

Table C.6: Missing data sensitivity analysis for Australia, lottery losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.84	-1.63	3.31	0.51	-2.73	3.74	2.04	0.73	3.34	1.57	1.41	1.73
Losses x 10^3	0.26	-0.98	1.50	1.69	-0.52	3.91	0.28	-0.97	1.53	0.21	-1.06	1.48
Losses^2 x 10^7	3.79	-4.01	11.60	60.29	5.84	114.75	3.62	-4.29	11.52	3.98	-4.03	12.00
Age	0.04	-0.03	0.10	0.05	-0.03	0.14	0.01	-0.05	0.06			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	0.06	-0.25	0.38	0.12	-0.26	0.50	0.08	-0.22	0.39			
Employment - Other	-1.61	-3.59	0.38	-2.39	-5.32	0.54						
Employment - Retired or pensioner	-0.33	-1.11	0.46	-0.01	-1.08	1.05						
Employment - Student	1.58	0.31	2.85	2.55	0.95	4.15						
Employment - Unemployed	0.68	-0.39	1.75	0.77	-0.63	2.18						
Employment - Working fulltime or parttime	-0.24	-0.91	0.43	-0.27	-1.09	0.55						
Education - School only	0.48	-1.31	2.27	0.41	-2.04	2.87						
Education - Technical	0.66	-1.15	2.48	0.85	-1.63	3.33						
Education - University	0.32	-1.48	2.12	0.45	-2.03	2.93						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital_status - Separated or divorced	0.66	0.12	1.19	0.66	0.02	1.30						
Marital_status - Single	0.46	0.01	0.91	0.58	0.03	1.13						
Marital_status - Widowed	0.08	-0.57	0.73	0.21	-0.65	1.07						
n	1029			722			973			976		
Missing	0%			30%			5%			5%		
adj. R^2	0.05			0.07			0.03			0.00		

Table C.7: Missing data sensitivity analysis for Australia, racing losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.71	-2.87	4.30	0.44	-3.53	4.41	2.93	1.34	4.52	1.74	1.54	1.94
Losses x 10^3	1.19	0.46	1.91	1.36	0.37	2.35	1.08	0.39	1.77	1.03	0.33	1.72
Losses^2 x 10^7	1.66	-0.55	3.88	1.59	-1.31	4.50	1.89	-0.25	4.03	2.12	-0.04	4.28
Age	0.00	-0.08	0.09	0.00	-0.11	0.11	-0.03	-0.11	0.04			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	-0.01	-0.45	0.44	0.09	-0.44	0.61	0.21	-0.21	0.63			
Employment - Other	0.67	-2.88	4.22	-2.54	-7.67	2.59						
Employment - Retired or pensioner	-0.22	-1.37	0.93	-0.21	-1.72	1.31						
Employment - Student	0.57	-1.07	2.22	1.39	-0.76	3.55						
Employment - Unemployed	1.95	0.21	3.70	-0.26	-2.51	1.98						
Employment - Working fulltime or parttime	0.02	-0.94	0.97	0.14	-1.11	1.39						
Education - School only	0.58	-2.15	3.31	0.66	-2.21	3.54						
Education - Technical	0.86	-1.91	3.63	1.07	-1.84	3.99						
Education - University	0.22	-2.54	2.99	0.27	-2.64	3.18						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital_status - Separated or divorced	0.66	-0.05	1.37	0.67	-0.16	1.50						
Marital_status - Single	0.54	-0.01	1.10	0.57	-0.11	1.24						
Marital_status - Widowed	0.24	-0.77	1.25	0.42	-0.88	1.73						
n	642			453			584			586		
Missing	0%			29%			9%			9%		
adj. R^2	0.09			0.08			0.07			0.05		

Table C.8: Missing data sensitivity analysis for Australia, sports betting losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	4.49	-2.02	11.00	4.64	-3.03	12.31	2.62	0.10	5.14	1.98	1.64	2.32
Losses x 10^3	-3.68	-11.87	4.51	-1.75	-12.64	9.15	-3.65	-12.05	4.76	-4.59	-13.04	3.85
$Losses^2$ x 10^7	17.79	-207.3	242.9	-36.89	-323.3	249.5	5.24	-226.1	236.5	20.34	-212.3	253.0
Age	0.01	-0.14	0.17	0.00	-0.22	0.21	-0.04	-0.16	0.08			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	0.99	0.21	1.77	1.10	0.07	2.12	0.69	-0.11	1.48			
Employment - Other	-4.76	-8.04	-1.48	-4.95	-9.29	-0.60						
Employment – Retired/pensioner	-3.72	-7.09	-0.36	-2.98	-7.54	1.57						
Employment - Student	-5.28	-9.83	-0.74	-5.79	-11.27	-0.31						
Employment - Unemployed	-4.99	-7.99	-1.98	-5.37	-9.36	-1.38						
Employment – Working	0.97	-3.99	5.93	1.24	-4.21	6.69						
Education - School only	1.35	-3.64	6.34	1.84	-3.66	7.34						
Education - Technical	1.08	-3.86	6.03	1.22	-4.19	6.62						
Education - University	0.00	0.00	0.00	0.00	0.00	0.00						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.60	-0.55	1.75	0.61	-0.78	2.00						
Marital_status – Separated/divorced	0.18	-0.75	1.11	0.28	-0.95	1.51						
Marital_status - Single	1.33	-1.09	3.76	1.20	-2.09	4.49						
Marital_status - Widowed	246			175			227			227		
n	0%			29%			8%			8%		
Missing	0.04			0.03			0.02			0.00		
adj. R^2	4.49	-2.02	11.00	4.64	-3.03	12.31	2.62	0.10	5.14	1.98	1.64	2.32

Table C.9: Missing data sensitivity analysis for Australia, table game losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	6.41	-0.02	12.85	6.47	-0.60	13.53	4.14	1.39	6.88	2.30	1.91	2.68
Losses x 10^3	1.38	-1.07	3.82	3.81	-1.04	8.65	1.36	-1.03	3.76	1.22	-1.16	3.60
$Losses^2$ x 10^7	-2.12	-8.88	4.65	-26.73	-67.13	13.68	-2.53	-9.43	4.37	-2.31	-9.18	4.57
Age	-0.15	-0.34	0.03	-0.26	-0.51	-0.02	-0.08	-0.22	0.06			
Age^2	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00			
Male	-0.20	-1.05	0.66	-0.17	-1.14	0.79	-0.21	-1.04	0.62			
Employment - Other	-2.77	-9.09	3.55	-1.51	-8.07	5.05						
Employment – Retired/pensioner	-3.20	-6.53	0.13	-2.74	-6.81	1.33						
Employment - Student	-2.49	-5.73	0.74	-0.59	-4.39	3.21						
Employment - Unemployed	1.33	-2.19	4.85	0.51	-3.57	4.59						
Employment – Working	-2.40	-5.20	0.39	-1.60	-4.86	1.66						
Education - School only	0.84	-2.96	4.64	1.36	-2.63	5.34						
Education - Technical	0.93	-2.98	4.84	1.77	-2.32	5.86						
Education - University	0.51	-3.30	4.33	1.04	-2.97	5.05						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital_status – Separated/divorced	1.25	-0.24	2.74	1.55	-0.21	3.31						
Marital_status - Single	0.07	-0.89	1.03	-0.30	-1.41	0.82						
Marital_status - Widowed	-0.85	-3.66	1.96	-1.59	-4.87	1.70						
n	239			169			225			225		
Missing	0%			29%			6%			6%		
adj. R^2	0.06			0.02			0.00			0.00		

Table C.10: Missing data sensitivity analysis for Finland, total gambling losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.81	0.39	1.23	0.73	0.28	1.18	1.18	0.88	1.47	0.16	0.12	0.20
Losses x 10^3	3.65	3.18	4.13	3.64	3.20	4.07	3.67	3.24	4.10	3.75	3.32	4.17
Losses^2 x 10^7	-4.80	-5.72	-3.88	-4.35	-5.24	-3.46	-4.80	-5.65	-3.96	-4.93	-5.77	-4.09
Age	-0.01	-0.03	0.01	0.00	-0.02	0.02	-0.04	-0.05	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.12	-0.19	-0.04	-0.14	-0.22	-0.05	-0.11	-0.19	-0.03			
Employment - Home duties	-0.08	-0.50	0.35	-0.11	-0.57	0.35						
Employment - Other	-0.16	-0.64	0.33	-0.13	-0.68	0.42						
Employment - Retired/pensioner	0.15	0.02	0.29	0.19	0.05	0.33						
Employment - Student	-0.18	-0.35	-0.01	-0.24	-0.43	-0.06						
Employment - Unemployed	0.10	-0.08	0.28	0.12	-0.08	0.31						
Education - School	0.12	0.04	0.20	0.11	0.02	0.19						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.20	0.07	0.33	0.21	0.07	0.35						
Marital status - Single	0.18	0.08	0.29	0.19	0.08	0.30						
Marital status - Widowed	0.22	0.02	0.41	0.25	0.04	0.46						
n	3450			3004			3396			3396		
Missing	0%			13%			2%			2%		
adj. R^2	0.16			0.17			0.12			0.10		

Table C.11: Missing data sensitivity analysis for Finland, EGM losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.09	-0.73	0.90	0.07	-0.79	0.93	0.77	0.23	1.31	0.13	0.04	0.22
Losses x 10^3	43.35	37.11	49.59	38.29	32.17	44.42	45.10	39.20	51.00	45.18	39.32	51.03
Losses^2 x 10^7	-822.6	-1114.8	-530.3	-627.6	-895.2	-360.0	-912.6	-1170.4	-654.8	-923.7	-1180.9	-666.4
Age	0.02	-0.01	0.06	0.03	-0.01	0.07	-0.02	-0.05	0.00			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.01	-0.18	0.16	-0.07	-0.25	0.10	0.02	-0.15	0.18			
Employment - Home duties	-0.28	-1.65	1.09	0.15	-1.23	1.53						
Employment - Other	-0.05	-0.97	0.87	0.11	-0.91	1.12						
Employment - Retired/pensioner	0.39	0.09	0.70	0.49	0.17	0.81						
Employment - Student	-0.26	-0.57	0.06	-0.28	-0.61	0.05						
Employment - Unemployed	0.19	-0.18	0.56	0.23	-0.17	0.62						
Education - School	0.17	0.00	0.34	0.17	-0.01	0.35						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.29	-0.04	0.62	0.32	-0.01	0.66						
Marital status - Single	0.27	0.06	0.48	0.31	0.09	0.52						
Marital status - Widowed	0.04	-0.54	0.62	0.09	-0.51	0.69						
n	1323			1156			1294			1294		
Missing	0%			13%			2%			2%		
adj. R^2	0.28			0.27			0.26			0.25		

Table C.12: Missing data sensitivity analysis for Finland, lotteries losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.98	0.52	1.45	1.10	0.61	1.59	1.14	0.79	1.50	0.17	0.12	0.22
Losses x 10^3	6.05	4.45	7.65	6.22	4.77	7.67	5.89	4.50	7.28	5.83	4.48	7.17
Losses^2 x 10^7	-45.22	-60.14	-30.30	-43.42	-57.77	-29.07	-44.28	-58.24	-30.33	-44.58	-58.31	-30.86
Age	-0.01	-0.03	0.01	-0.01	-0.04	0.01	-0.03	-0.05	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.19	-0.27	-0.10	-0.19	-0.28	-0.10	-0.15	-0.23	-0.07			
Employment - Home duties	-0.04	-0.49	0.42	-0.06	-0.54	0.43						
Employment - Other	0.02	-0.49	0.54	0.07	-0.53	0.66						
Employment - Retired/pensioner	0.09	-0.05	0.24	0.09	-0.07	0.24						
Employment - Student	-0.30	-0.50	-0.11	-0.30	-0.51	-0.10						
Employment - Unemployed	0.04	-0.15	0.23	0.09	-0.11	0.29						
Education - School	0.10	0.01	0.18	0.12	0.03	0.21						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.23	0.10	0.37	0.19	0.05	0.33						
Marital status - Single	0.19	0.07	0.30	0.21	0.09	0.33						
Marital status - Widowed	0.21	0.00	0.41	0.23	0.02	0.45						
n	3064			2700			3010			3010		
Missing	0%			12%			2%			2%		
adj. R^2	0.10			0.12			0.04			0.02		

Table C.13: Missing data sensitivity analysis for Finland, racing losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	2.61	-0.07	5.28	2.12	-0.63	4.87	1.90	-0.13	3.94	0.76	0.54	0.99
Losses x 10^3	-0.56	-3.16	2.03	-1.11	-3.81	1.59	-0.38	-2.98	2.21	-0.24	-2.82	2.33
$Losses^2$ x 10^7	4.45	-2.92	11.83	6.65	-1.31	14.61	3.68	-3.70	11.06	3.43	-3.92	10.77
Age	-0.02	-0.13	0.10	-0.01	-0.13	0.11	-0.03	-0.12	0.06			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.64	-1.11	-0.16	-0.57	-1.05	-0.09	-0.42	-0.86	0.01			
Employment - Home duties	2.65	0.38	4.91	5.81	2.67	8.96						
Employment - Other	0.90	-1.58	3.38	3.09	-0.18	6.37						
Employment - Retired/pensioner	-0.60	-1.34	0.14	-0.63	-1.39	0.12						
Employment - Student	-1.07	-2.39	0.26	-0.83	-2.15	0.48						
Employment - Unemployed	-0.18	-1.11	0.75	-0.06	-0.98	0.87						
Education - School	0.14	-0.33	0.61	0.18	-0.32	0.67						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.86	0.21	1.52	0.66	-0.02	1.35						
Marital status - Single	-0.11	-0.70	0.49	-0.14	-0.74	0.47						
Marital status - Widowed	0.39	-0.60	1.37	0.29	-0.73	1.30						
n	241			215			234			234		
Missing	0%			11%			3%			3%		
adj. R^2	0.15			0.17			0.08			0.06		

Table C.14: Missing data sensitivity analysis for Finland, sports betting losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	2.29	0.68	3.90	2.44	0.69	4.19	2.30	1.21	3.38	0.53	0.35	0.70
Losses x 10^3	9.95	6.90	13.01	9.66	6.53	12.78	10.11	7.18	13.04	9.93	7.00	12.87
$Losses^2$ x 10^7	-50.62	-70.66	-30.57	-49.20	-69.76	-28.63	-52.18	-71.69	-32.67	-51.88	-71.50	-32.26
Age	-0.06	-0.13	0.01	-0.06	-0.14	0.01	-0.07	-0.12	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.06	-0.47	0.35	-0.05	-0.49	0.39	-0.03	-0.43	0.37			
Employment - Home duties	4.45	1.13	7.76	4.39	0.97	7.81						
Employment - Other	-0.86	-2.82	1.11	-0.79	-3.28	1.70						
Employment - Retired/pensioner	0.07	-0.53	0.67	0.10	-0.54	0.73						
Employment - Student	-0.70	-1.40	-0.01	-0.73	-1.52	0.06						
Employment - Unemployed	0.18	-0.47	0.83	0.20	-0.53	0.92						
Education - School	0.39	0.06	0.71	0.42	0.06	0.77						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.46	-0.11	1.04	0.30	-0.32	0.91						
Marital status - Single	0.24	-0.17	0.65	0.23	-0.21	0.66						
Marital status - Widowed	0.34	-0.68	1.37	0.25	-0.81	1.31						
n	488			435			479			479		
Missing	0%			11%			2%			2%		
adj. R^2	0.15			0.14			0.11			0.09		

Table C.15: Missing data sensitivity analysis for Finland, table games losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.63	-3.18	4.44	1.02	-3.01	5.06	0.84	-1.71	3.38	0.59	0.24	0.94
Losses x 10^3	30.62	12.77	48.47	30.09	11.38	48.80	33.35	16.60	50.11	33.69	17.24	50.15
$Losses^2$ x 10^7	-226.05	-1036.92	584.82	-225.87	-1067.41	615.68	-386.02	-1130.12	358.07	-383.77	-1117.68	350.14
Age	0.05	-0.14	0.23	0.05	-0.15	0.25	0.00	-0.15	0.14			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	0.11	-0.75	0.97	0.09	-0.79	0.98	0.17	-0.64	0.99			
Employment - Home duties	1.31	-3.09	5.72	1.13	-3.40	5.65						
Employment - Other	-1.95	-4.23	0.32	-2.38	-4.91	0.15						
Employment - Retired/pensioner	1.14	-0.51	2.79	1.89	-0.17	3.94						
Employment - Student	-0.64	-1.89	0.60	-0.76	-2.06	0.53						
Employment - Unemployed	-0.08	-1.34	1.18	-0.30	-1.69	1.09						
Education - School	0.46	-0.24	1.16	0.52	-0.22	1.26						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	1.01	-1.06	3.08	0.02	-2.10	2.15						
Marital status - Single	0.12	-0.62	0.86	0.07	-0.71	0.85						
Marital status - Widowed	191			172			185			185		
n	0%			10%			3%			3%		
Missing	0.19			0.19			0.18			0.19		
adj. R^2	0.63	-3.18	4.44	1.02	-3.01	5.06	0.84	-1.71	3.38	0.59	0.24	0.94

Table C.16: Missing data sensitivity analysis for Canada, total gambling losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	1.50	0.89	2.10	1.12	0.37	1.87	1.29	0.85	1.72	0.12	0.05	0.18
Losses x 10^3	2.38	1.95	2.81	2.01	1.54	2.49	2.31	1.92	2.70	2.55	2.15	2.95
Losses^2 x 10^7	-4.86	-6.21	-3.52	-3.88	-5.33	-2.43	-4.71	-6.06	-3.37	-5.31	-6.70	-3.92
Age	-0.04	-0.06	-0.02	-0.03	-0.06	0.01	-0.04	-0.06	-0.03			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.16	-0.28	-0.04	-0.07	-0.21	0.07	-0.13	-0.24	-0.02			
Employment - Other	0.07	-0.50	0.63	0.11	-0.48	0.71						
Employment - Retired or pensioner	-0.04	-0.33	0.25	-0.06	-0.48	0.35						
Employment - Student	-0.14	-0.50	0.22	-0.22	-0.74	0.30						
Employment - Unemployed	0.09	-0.30	0.48	0.17	-0.31	0.65						
Employment - Employed	-0.09	-0.31	0.12	-0.03	-0.35	0.28						
Education - Technical	-0.16	-0.29	-0.02	-0.13	-0.30	0.04						
Education - University	-0.10	-0.24	0.04	-0.08	-0.25	0.09						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - separated/divorced	0.13	-0.04	0.30	0.16	-0.05	0.38						
Marital status - Single	0.02	-0.14	0.18	0.07	-0.12	0.27						
Marital status - Widowed	0.00	-0.26	0.26	-0.09	-0.52	0.33						
n	2645			1259			2036			2075		
Missing	0%			52%			23%			22%		
adj. R^2	0.10			0.08			0.09			0.08		

Table C.17: Missing data sensitivity analysis for Canada, EGM losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	2.53	1.25	3.81	1.67	0.06	3.29	2.02	1.18	2.85	0.40	0.25	0.55
Losses x 10^3	3.42	2.11	4.72	3.25	1.92	4.58	3.13	1.99	4.28	3.56	2.40	4.72
$Losses^2$ x 10^7	-8.71	-13.90	-3.52	-8.10	-12.39	-3.82	-7.92	-11.81	-4.02	-9.27	-13.29	-5.25
Age	-0.05	-0.10	0.00	-0.03	-0.10	0.04	-0.06	-0.10	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.31	-0.58	-0.05	-0.03	-0.35	0.29	-0.28	-0.52	-0.03			
Employment - Other	-0.30	-1.47	0.86	0.24	-1.22	1.70						
Employment - Retired or pensioner	0.27	-0.50	1.05	0.12	-0.85	1.08						
Employment - Student	-0.60	-1.48	0.28	-0.29	-1.39	0.82						
Employment - Unemployed	0.68	-0.25	1.62	0.53	-0.69	1.76						
Employment - Employed	-0.09	-0.69	0.51	0.01	-0.69	0.71						
Education - Technical	-0.27	-0.58	0.05	-0.18	-0.56	0.20						
Education - University	-0.15	-0.46	0.17	-0.09	-0.47	0.28						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - separated/divorced	0.44	0.01	0.87	0.28	-0.23	0.80						
Marital status - Single	-0.01	-0.37	0.35	0.08	-0.35	0.51						
Marital status - Widowed	-0.31	-1.03	0.42	-0.34	-1.24	0.56						
n	800			455			739			755		
Missing	0%			43%			8%			6%		
adj. R^2	0.09			0.06			0.06			0.05		

Table C.18: Missing data sensitivity analysis for Canada, lottery losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	1.89	1.15	2.62	1.63	0.71	2.54	1.74	1.19	2.29	0.22	0.12	0.32
Losses x 10^3	9.53	4.43	14.63	12.21	5.57	18.85	10.89	5.86	15.92	10.77	5.61	15.92
$Losses^2$ x 10^7	-218.94	-477.97	40.10	-450.24	-842.31	-58.17	-280.71	-539.42	-21.99	-254.73	-523.56	14.11
Age	-0.05	-0.07	-0.02	-0.04	-0.08	0.00	-0.06	-0.08	-0.03			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.19	-0.32	-0.06	-0.06	-0.23	0.11	-0.13	-0.26	0.00			
Employment - Other	-0.05	-0.55	0.46	0.06	-0.63	0.74						
Employment – Retired/pensioner	-0.02	-0.34	0.31	-0.01	-0.50	0.48						
Employment - Student	-0.05	-0.55	0.44	-0.27	-0.95	0.42						
Employment - Unemployed	0.14	-0.29	0.57	0.30	-0.28	0.87						
Employment - Employed	-0.06	-0.31	0.19	-0.03	-0.39	0.34						
Education - Technical	-0.21	-0.36	-0.05	-0.23	-0.42	-0.03						
Education - University	-0.11	-0.28	0.06	-0.09	-0.29	0.12						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - separated/divorced	0.18	-0.04	0.41	0.24	0.00	0.49						
Marital status - Single	0.02	-0.18	0.22	0.06	-0.16	0.29						
Marital status - Widowed	0.00	-0.31	0.30	-0.07	-0.55	0.42						
n	2162			1073			1717			1751		
Missing	0%			50%			21%			19%		
adj. R^2	0.05			0.04			0.04			0.02		

Table C.19: Missing data sensitivity analysis for Canada, racing losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	3.39	0.53	6.26	1.05	-4.08	6.18	3.12	1.42	4.83	0.18	-0.26	0.62
Losses x 10^3	17.37	5.33	29.41	7.94	-6.79	22.67	15.88	5.04	26.72	17.42	6.82	28.01
Losses^2 x 10^7	-371.78	-634.23	-109.33	-200.89	-516.84	115.06	-340.02	-578.77	-101.27	-367.91	-607.82	-128.00
Age	-0.10	-0.19	-0.02	0.00	-0.23	0.24	-0.11	-0.18	-0.04			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	-0.31	-1.01	0.39	-0.83	-1.73	0.07	-0.34	-0.95	0.28			
Employment - Other	-0.70	-4.15	2.75	-0.70	-3.94	2.55						
Employment - Retired or pensioner	-0.74	-2.83	1.36	-1.12	-4.47	2.23						
Employment - Student	0.24	-2.02	2.50	1.29	-1.09	3.66						
Employment - Unemployed	-1.56	-3.90	0.78	-1.57	-4.93	1.80						
Employment - Employed	-0.62	-2.16	0.92	-0.90	-2.41	0.61						
Education - Technical	-0.27	-1.11	0.56	0.31	-0.72	1.34						
Education - University	-0.18	-0.96	0.59	0.89	-0.12	1.90						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - separated/divorced	-0.22	-1.29	0.84	0.15	-1.33	1.63						
Marital status - Single	-0.03	-1.06	1.00	-0.42	-1.68	0.84						
Marital status - Widowed	0.58	-1.92	3.08	-0.75	-3.47	1.96						
n	117			68			113			115		
Missing	0%			42%			3%			2%		
adj. R^2	0.09			0.01			0.14			0.07		

Table C.20: Missing data sensitivity analysis for Canada, sports betting losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	2.18	-0.05	4.41	1.79	-0.94	4.51	1.83	0.74	2.91	0.46	0.23	0.69
Losses x 10^3	7.76	1.02	14.49	7.85	0.86	14.84	8.88	3.37	14.38	10.14	3.84	16.44
Losses^2 x 10^7	-54.6	-129.5	20.3	-55.7	-131.7	20.2	-70.9	-131.5	-10.4	-86.1	-155.8	-16.4
Age	-0.06	-0.13	0.01	-0.06	-0.18	0.07	-0.06	-0.10	-0.01			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	0.01	-0.45	0.47	0.12	-0.39	0.62	-0.12	-0.49	0.25			
Employment - Other	-0.29	-4.05	3.47	-0.13	-3.51	3.25						
Employment - Retired or pensioner	-0.14	-1.68	1.39	0.13	-1.66	1.92						
Employment - Student	-0.17	-1.68	1.35	0.25	-1.51	2.01						
Employment - Unemployed	-0.02	-1.82	1.78	1.01	-1.25	3.28						
Employment - Employed	0.58	-0.60	1.76	0.42	-0.96	1.81						
Education - Technical	-0.74	-1.24	-0.24	-0.52	-1.06	0.03						
Education - University	-0.81	-1.29	-0.33	-0.79	-1.30	-0.28						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - separated/divorced	0.82	0.14	1.51	1.01	0.33	1.68						
Marital status - Single	0.26	-0.33	0.85	0.51	-0.09	1.11						
Marital status - Widowed	-0.34	-1.84	1.16	-0.15	-2.36	2.07						
n	356			216			320			329		
Missing	0%			39%			10%			8%		
adj. R^2	0.09			0.10			0.06			0.03		

Table C.21: Missing data sensitivity analysis for Canada, table games losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	4.73	0.83	8.63	1.98	-2.59	6.55	3.26	1.43	5.09	1.11	0.65	1.57
Losses x 10^3	2.26	-0.81	5.34	0.30	-2.90	3.50	2.55	-0.13	5.23	1.73	-1.29	4.76
$Losses^2$ x 10^7	-14.49	-34.80	5.82	-1.43	-20.95	18.09	-15.68	-33.66	2.29	-10.99	-31.54	9.57
Age	-0.10	-0.23	0.03	0.00	-0.18	0.18	-0.10	-0.18	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Female	0.18	-0.62	0.99	0.91	0.03	1.79	0.13	-0.55	0.80			
Employment - Other	1.08	-3.05	5.20	-0.41	-4.27	3.45						
Employment - Retired or pensioner	1.07	-1.67	3.81	0.36	-2.57	3.30						
Employment - Student	0.71	-1.96	3.38	3.74	-0.23	7.71						
Employment - Unemployed	2.93	0.10	5.75	2.85	-1.03	6.73						
Employment - Employed	0.87	-1.24	2.98	0.52	-2.20	3.23						
Education - Technical	-1.00	-1.91	-0.09	-0.36	-1.29	0.58						
Education - University	-1.12	-1.94	-0.31	-0.24	-1.19	0.71						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
$Income^2$	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - separated/divorced	-0.30	-1.40	0.80	0.18	-0.99	1.35						
Marital status - Single	-0.66	-1.74	0.41	-0.67	-1.80	0.45						
Marital status - Widowed	-0.76	-3.73	2.20	-1.22	-4.69	2.25						
n	202			118			188			191		
Missing	0%			42%			7%			5%		
adj. R^2	0.08			0.04			0.03			0.00		

Table C.22: Missing data sensitivity analysis for Norway, total gambling losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.64	0.38	0.91	0.54	0.27	0.80	0.79	0.61	0.97	0.01	-0.02	0.03
Losses x 10^3	1.75	1.53	1.98	1.64	1.41	1.87	1.76	1.54	1.99	1.85	1.63	2.08
Losses^2 x 10^7	-2.98	-3.57	-2.40	-2.59	-3.16	-2.01	-3.00	-3.58	-2.41	-3.16	-3.75	-2.57
Age	-0.03	-0.04	-0.02	-0.02	-0.03	-0.01	-0.03	-0.04	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	0.00	-0.04	0.05	0.00	-0.05	0.04	0.01	-0.03	0.05			
Employment - Employed	-0.03	-0.09	0.04	-0.01	-0.07	0.06						
Education - Post-school	-0.03	-0.08	0.02	-0.02	-0.07	0.03						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.10	0.01	0.18	0.11	0.03	0.20						
Marital status - Single	0.11	0.03	0.18	0.07	-0.01	0.14						
Marital status - Widowed	-0.02	-0.15	0.11	-0.05	-0.19	0.08						
n	2184			1875			2184			2184		
Missing	0%			14%			0%			0%		
adj. R^2	0.17			0.15			0.16			0.13		

Table C.23: Missing data sensitivity analysis for Norway, EGM losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.81	-0.79	2.41	0.49	-1.32	2.29	0.76	-0.14	1.66	0.33	0.17	0.50
Losses x 10^3	5.99	4.48	7.50	5.49	3.66	7.32	5.89	4.36	7.41	5.93	4.43	7.43
Losses^2 x 10^7	-11.10	-14.96	-7.23	-9.94	-14.52	-5.36	-10.89	-14.75	-7.02	-10.87	-14.65	-7.09
Age	0.00	-0.07	0.07	0.02	-0.06	0.10	-0.03	-0.08	0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	-0.02	-0.41	0.37	-0.03	-0.51	0.46	0.04	-0.34	0.42			
Employment - Employed	-0.10	-0.48	0.29	-0.05	-0.52	0.41						
Education - Post-school	-0.11	-0.54	0.33	-0.10	-0.56	0.36						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.58	-0.19	1.35	0.56	-0.24	1.35						
Marital status - Single	0.28	-0.18	0.74	0.29	-0.23	0.80						
Marital status - Widowed	-1.07	-2.83	0.69	-0.90	-2.73	0.93						
n	234			180			234			234		
Missing	0%			23%			0%			0%		
adj. R^2	0.29			0.24			0.26			0.26		

Table C.24: Missing data sensitivity analysis for Norway, lotteries losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	0.64	0.39	0.89	0.37	0.10	0.63	0.73	0.53	0.94	0.02	0.00	0.05
Losses x 10^3	1.77	1.19	2.35	3.10	2.25	3.95	1.81	1.23	2.39	1.81	1.22	2.39
Losses^2 x 10^7	-6.99	-9.47	-4.50	-30.78	-44.95	-16.62	-7.09	-9.57	-4.61	-7.07	-9.58	-4.57
Age	-0.03	-0.04	-0.02	-0.01	-0.02	0.00	-0.03	-0.04	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	0.05	0.01	0.09	0.04	0.00	0.08	0.05	0.02	0.09			
Employment - Employed	-0.01	-0.07	0.05	-0.02	-0.08	0.05						
Education - Post-school	-0.03	-0.08	0.02	-0.01	-0.06	0.04						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.10	0.02	0.18	0.06	-0.02	0.14						
Marital status - Single	0.14	0.06	0.21	0.09	0.02	0.17						
Marital status - Widowed	0.01	-0.11	0.13	-0.01	-0.14	0.12						
n	2232			1943			2232			2232		
Missing	0%			13%			0%			0%		
adj. R^2	0.05			0.04			0.04			0.02		

Table C.25 Missing data sensitivity analysis for Norway, racing losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	4.05	1.17	6.92	2.82	-0.15	5.79	3.79	1.87	5.71	0.47	0.18	0.76
Losses x 10^3	1.03	-0.57	2.63	0.77	-0.89	2.43	0.46	-1.19	2.11	0.45	-1.25	2.16
Losses^2 x 10^7	-4.12	-9.52	1.28	-3.37	-8.92	2.19	-2.21	-7.70	3.27	-1.74	-7.45	3.97
Age	-0.12	-0.23	-0.02	-0.04	-0.16	0.08	-0.13	-0.22	-0.05			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	-0.34	-0.94	0.27	-0.33	-0.96	0.31	-0.19	-0.78	0.41			
Employment - Employed	0.47	-0.36	1.30	0.11	-0.86	1.08						
Education - Post-school	-0.29	-0.89	0.31	-0.14	-0.73	0.44						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	2.56	1.05	4.07	2.25	0.81	3.69						
Marital status - Single	0.41	-0.35	1.17	0.45	-0.34	1.23						
Marital status - Widowed	-0.60	-2.14	0.95	-0.75	-2.43	0.92						
n	115			101			115			115		
Missing	0%			12%			0%			0%		
adj. R^2	0.20			0.12			0.09			-0.01		

Table C.26: Missing data sensitivity analysis for Norway, sports betting losses

	Model 1 - Full model, multiple imputation			Model 2 - Full model, complete cases			Model 3 - Reduced model, complete cases			Model 4 - Bivariate model, complete cases		
	Estimate	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**		**Estimate**	**95% c.i.**	
Constant	1.27	0.43	2.10	1.22	0.36	2.08	1.19	0.62	1.75	0.06	-0.03	0.15
Losses x 10^3	4.05	2.89	5.21	4.51	3.37	5.66	3.91	2.76	5.06	4.17	3.03	5.31
Losses^2 x 10^7	-25.58	-35.19	-15.97	-28.35	-37.71	-18.98	-24.82	-34.31	-15.33	-25.62	-35.18	-16.05
Age	-0.03	-0.07	0.00	-0.02	-0.06	0.02	-0.05	-0.07	-0.02			
Age^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Male	-0.10	-0.31	0.12	-0.17	-0.39	0.05	-0.07	-0.28	0.13			
Employment - Employed	-0.06	-0.29	0.17	-0.09	-0.33	0.15						
Education - Post-school	-0.12	-0.30	0.05	-0.09	-0.26	0.09						
Income	0.00	0.00	0.00	0.00	0.00	0.00						
Income^2	0.00	0.00	0.00	0.00	0.00	0.00						
Marital status - Separated/divorced	0.25	-0.13	0.63	0.16	-0.22	0.54						
Marital status - Single	0.07	-0.15	0.30	0.01	-0.22	0.24						
Marital status - Widowed	-0.33	-0.88	0.22	-0.52	-1.14	0.10						
n	460			402			460			460		
Missing	0%			13%			0%			0%		
adj. R^2	0.15			0.16			0.14			0.10		

Appendix D: Supplementary figures and tables for Chapter 4

Table D.1: Predictors of the prevalence of gambling-harm in EGM venues, weighted by raw respondent count and weighted by EGM count.

	Coefficient estimate (95% confidence interval)	
	Weighted by raw respondent count	Weighted by venue EGM count
Intercept	-3.33 (-4.12, -2.53) ***	-3.18 (-3.99, -2.37) ***
Monthly expenditure per adult, 100s AUD	0.73 (0.28, 1.18) **	0.66 (0.19, 1.14) **
Venue type		
Casino	0.00 (ref. group)	0.00 (ref. group)
Club	0.65 (0.20, 1.09) **	0.59 (0.11, 1.08) *
Hotel	0.21 (-0.15, 0.57)	0.14 (-0.20, 0.48)
Number of EGMs, 10s	0.01 (0.01, 0.02) ***	0.01 (0.01, 0.01) ***

Notes: Deviance explained was 40% and 69% for the model weighted by respondent count and EGM count, respectively. *N* venues = 62 in all models. . * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Coefficients are expressed on the logit scale. *P* values and confidence intervals have been corrected for heteroskedasticity .

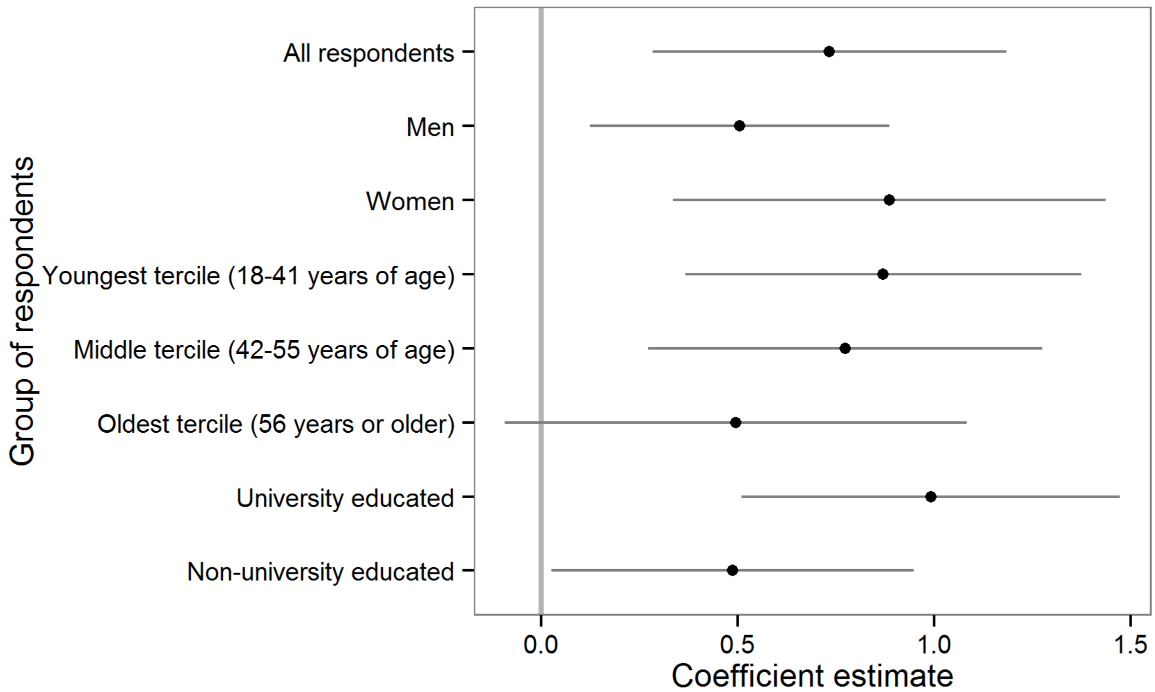

Figure D.1: Coefficient estimates (on the logit scale) and heteroskedasticity corrected 95% confidence intervals for the association between mean monthly EGM expenditure per adult and rates of gambling-related harm among important subpopulations in our study. Coefficients were estimated by multivariate binomial regression. For the details of the full models, see Table D.2.

Table D.2: Predictors of the prevalence of gambling-harm in EGM venues for population subgroups.

	All respondents	Women	Men	Youngest tercile (18-41 years of age)
Intercept	-3.33 (-4.12, -2.53) ***	-3.84 (-4.79, -2.89) ***	-2.55 (-3.24, -1.85) ***	-3.55 (-4.42, -2.68) ***
Monthly expenditure per adult, 100s AUD	0.73 (0.28, 1.18) **	0.89 (0.34, 1.44) **	0.51 (0.12, 0.89) **	0.87 (0.37, 1.37) ***
Venue type				
Casino	0.0 (ref. group)	0.0 (ref. group)	0.0 (ref. group)	0.0 (ref. group)
Club	0.65 (0.20, 1.09) **	0.85 (0.35, 1.35) ***	0.22 (-0.18, 0.62)	0.89 (0.34, 1.43) **
Hotel	0.21 (-0.15, 0.57)	0.14 (-0.27, 0.55)	0.09 (-0.26, 0.44)	0.36 (-0.08, 0.80)
Number of EGMs, 10s	0.01 (0.01, 0.02) ***	0.01 (0.01, 0.02) ***	0.01 (0.00, 0.01) ***	0.01 (0.01, 0.02) ***
Deviance explained	40%	47%	29%	24%
n respondents	4950	2982	1887	1750

	Middle tercile (42-55 years of age)	Oldest tercile (56 years or older)	University education	Non-university education
Intercept	-4.45 (-5.33, -3.56) ***	-3.79 (-4.81, -2.77) ***	-4.16 (-5.00, -3.31) ***	-2.74 (-3.54, -1.93) ***
Monthly expenditure per adult, 100s AUD	0.77 (0.27, 1.27) **	0.50 (-0.09, 1.08)	0.99 (0.51, 1.47) ***	0.49 (0.03, 0.95) *
Venue type				
Casino	0.0 (ref. group)	0.0 (ref. group)	0.0 (ref. group)	0.0 (ref. group)
Club	1.73 (1.23, 2.23) ***	1.10 (0.47, 1.73) ***	0.88 (0.42, 1.34) ***	0.43 (-0.10, 0.97)
Hotel	1.20 (0.78, 1.62) ***	1.46 (0.93, 1.98) ***	0.30 (-0.08, 0.69)	0.18 (-0.21, 0.58)
Number of EGMs, 10s	0.03 (0.02, 0.03) ***	0.02 (0.02, 0.03) ***	0.02 (0.01, 0.02) ***	0.01 (0.01, 0.01) ***
Deviance explained	30%	24%	49%	25%
n respondents	1566	1485	2211	2731

Notes: Coefficients are expressed on the logit scale. *P* values and Wald-type confidence intervals have been calculated using heteroskedasticity correction. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *N* venues = 62 in all models. *n* respondents in each model is the number of respondents in each subpopulation who reported visiting an EGM venue in the last month and completed the PGSI. The raw number of respondents per venue is used as weights in the binomial regression model. *P* values reported in this table should be treated with caution as no adjustment has been made for multiple comparisons.

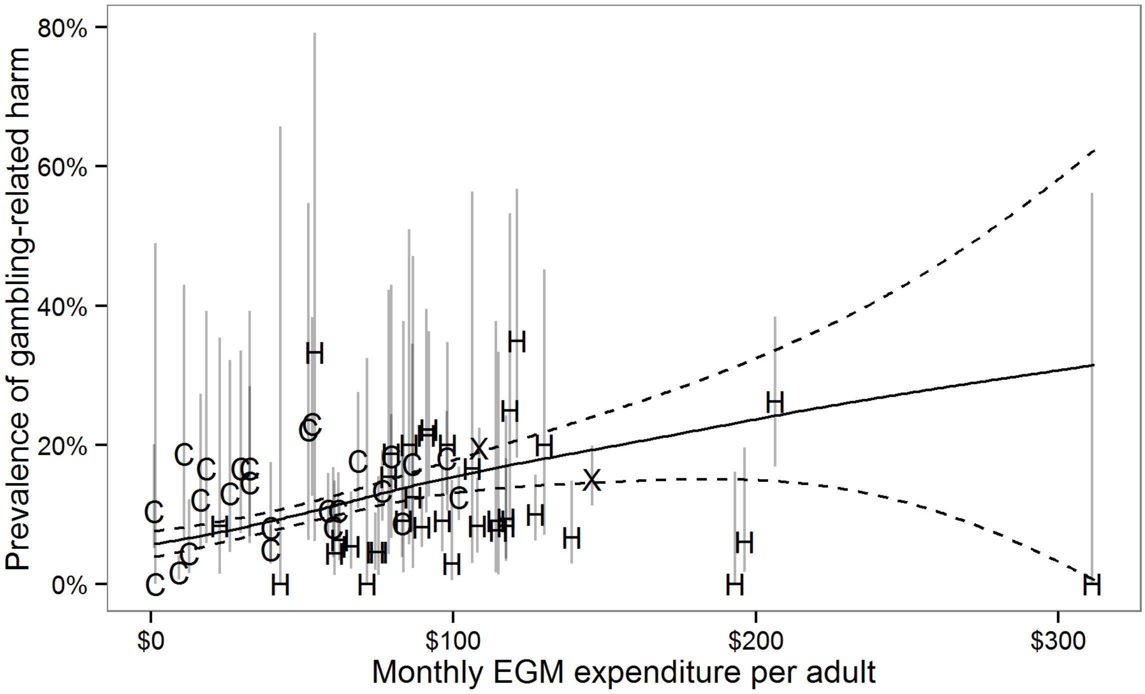

Figure D.2: Predicted prevalence of gambling-related harm for a hypothetical club with the median number of EGMs (22), estimated using a semi-parametric spline in a generalised additive model. The solid black line shows the fitted regression line, and the dashed black lines outline the 95% confidence bounds. Points indicate actual venues in the study. Symbols X, C and H indicate venues of type casino, club and hotel, respectively. The intersecting vertical grey lines showing the 95% confidence interval for the prevalence of gambling-related harm at that venue, calculated using Wilson's method. Wilson's confidence intervals are asymmetric except when P = 0.5.

Appendix E: Supplementary figures and tables for Chapter 5

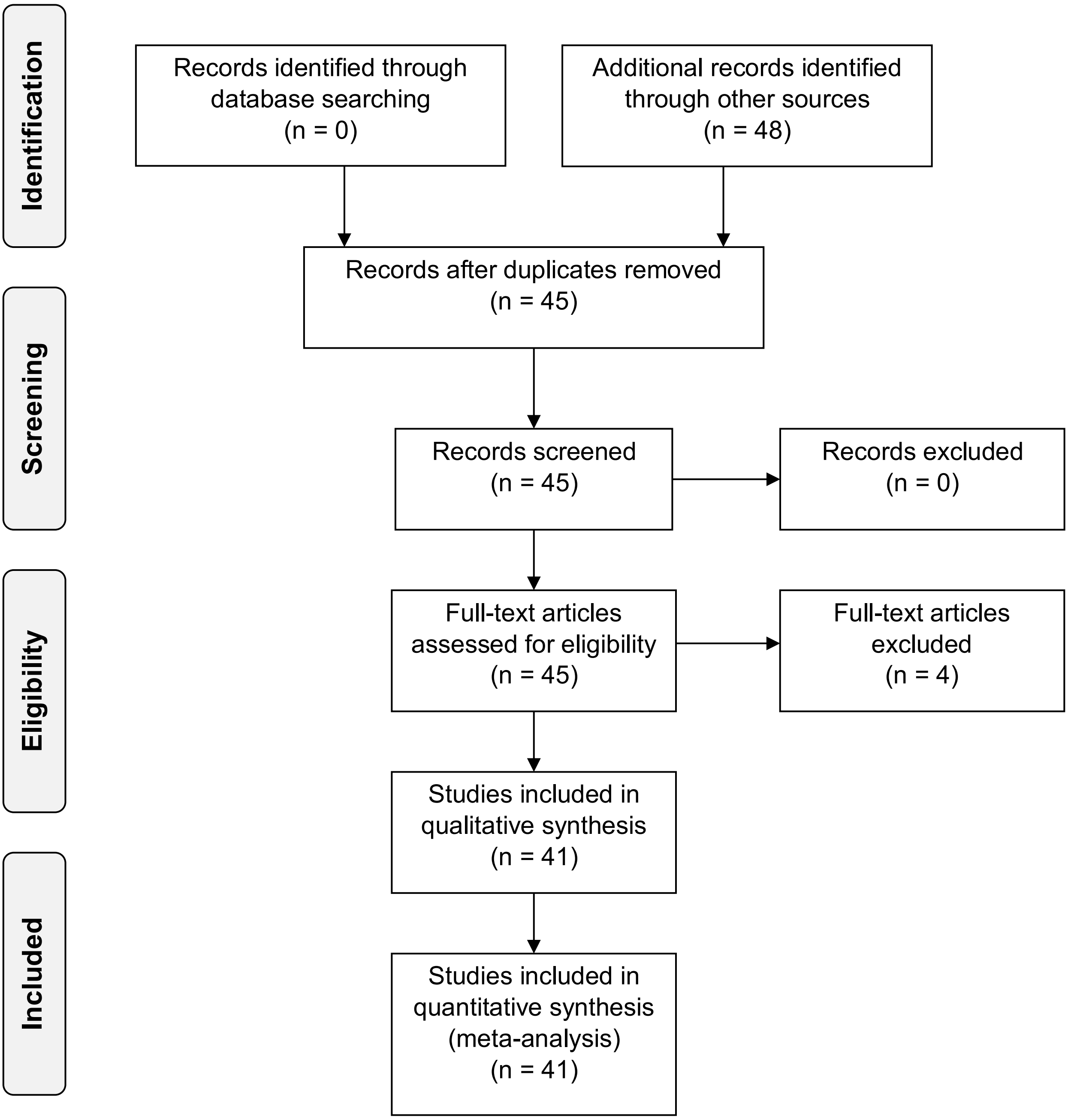

Figure E.1: Prevalence study records identified, included and excluded.

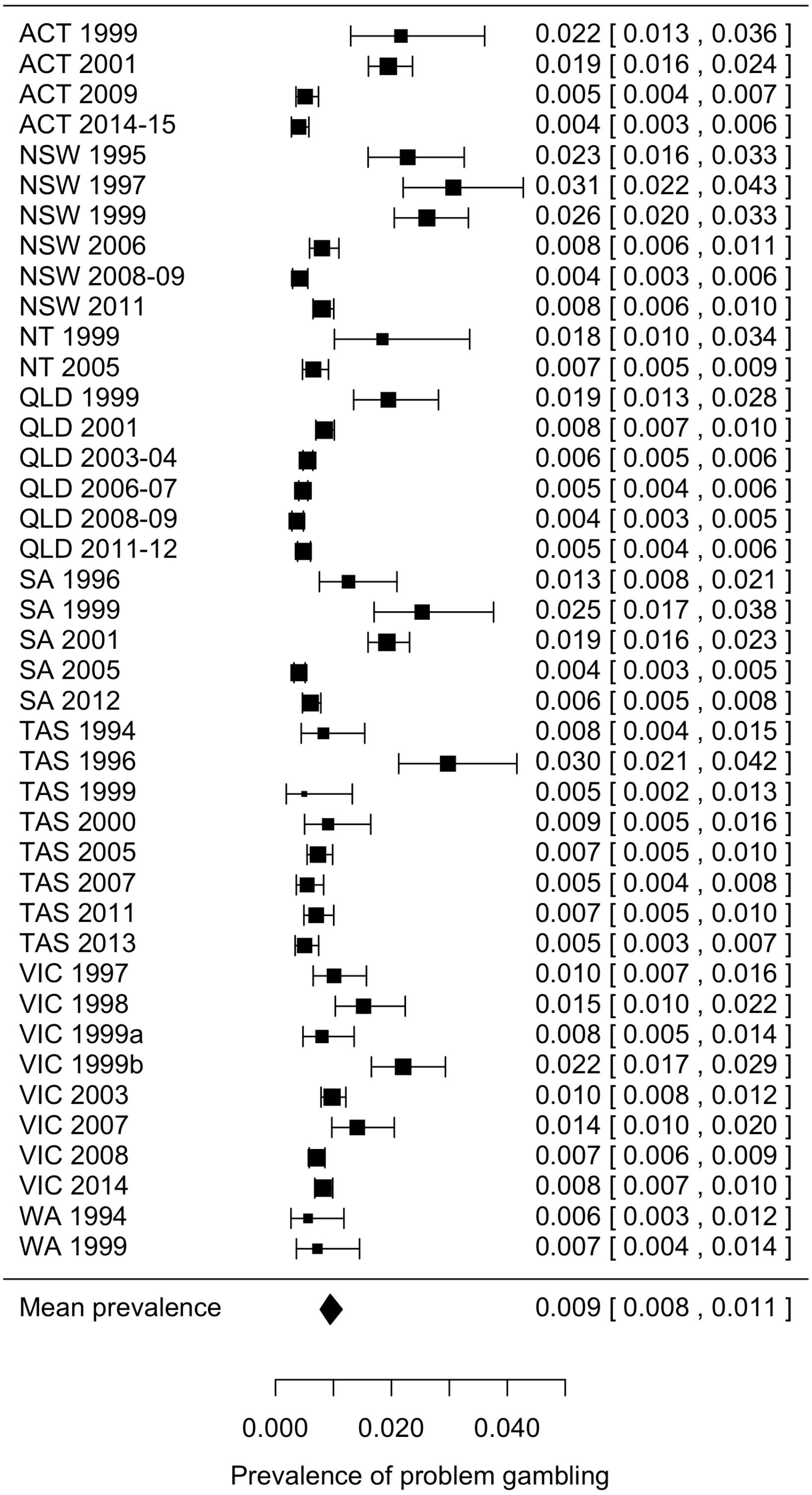

Figure E.2: Problem gambling prevalence estimates in all individual studies (*n* = 41) and mean prevalence estimated by random effects meta-analysis.

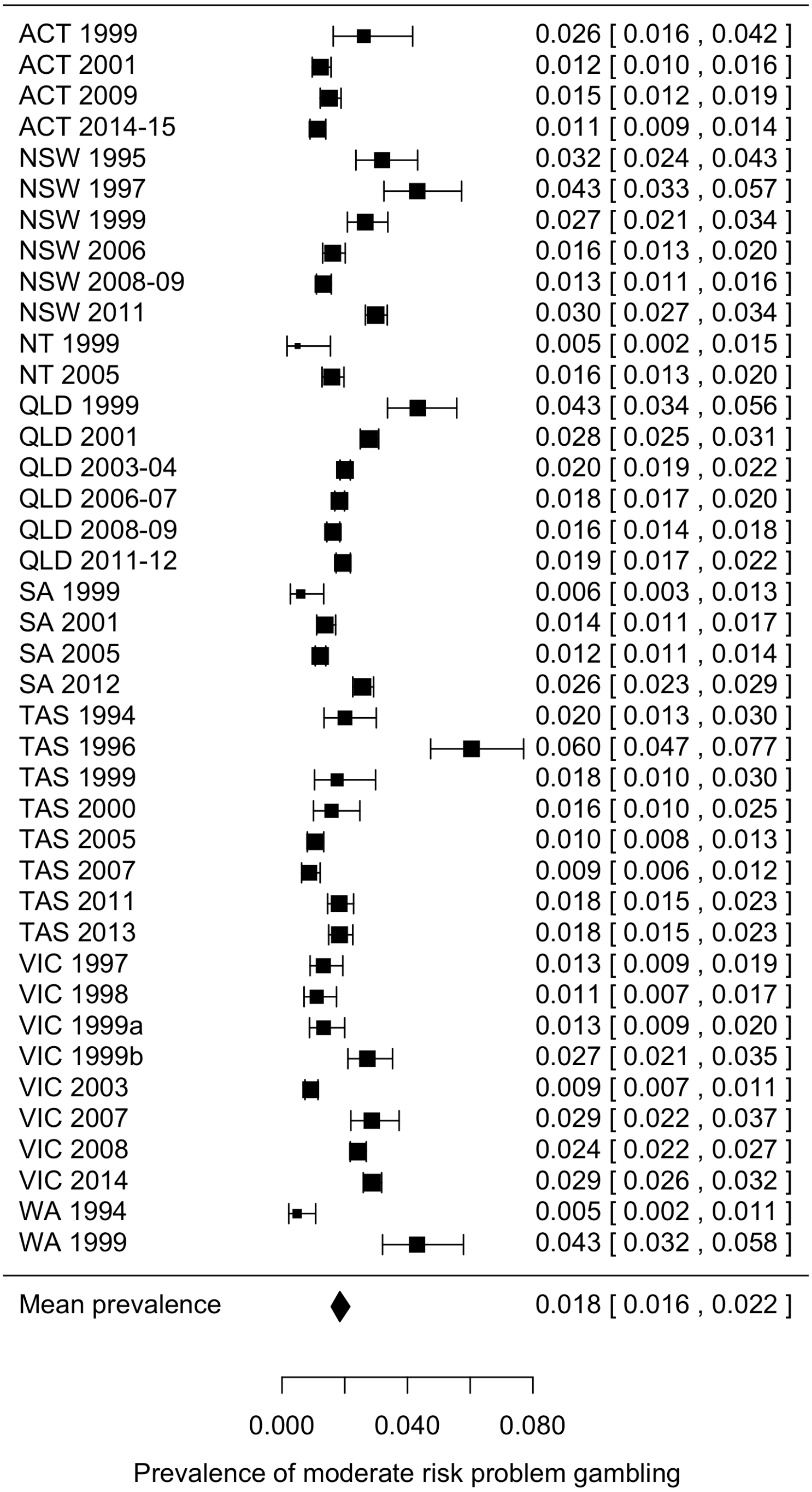

Figure E.3: Moderate risk problem gambling prevalence estimates in all individual studies (*n* = 40) and mean prevalence estimated by random effects meta-analysis.

Table E.1: Full bibliographic details for each eligible study

State or territory	Year	Full bibliography
ACT	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E
ACT	2001	Tremayne K, Masterman-Smith H, McMillen J. Survey of the nature and extent of gambling and problem gambling in the ACT [Internet]. Sydney: Australian Institute for Gambling Research, University of Western Sydney; 2001. Available from: http://www.gamblingandracing.act.gov.au/__data/assets/pdf_file/0009/745065/Survey-of-Problem-Gambling-in-the-ACT.pdf archived at http://www.webcitation.org/6k8jsBDpr
ACT	2009	Davidson T, Rodgers B. 2009 Survey of the nature and extent of gambling, and problem gambling in the Australian Capital Territory [Internet]. Canberra: The Centre for Gambling Research, The Australian National University; 2010. Available from: http://www.problemgambling.act.gov.au/Recent%20Research/ACT%20Gambling%20Prevalence%20Study.pdf archived at http://www.webcitation.org/6k8kG4vKu
ACT	2014-15	Davidson T, Rodgers B, Taylor-Rodgers E, Suomi A, Lucas N. 2014 survey on gambling, health and wellbeing in the ACT [Internet]. Canberra: Centre for Gambling Research, The Australian National University; 2015. Available from: http://sociology.cass.anu.edu.au/sites/default/files/2014%20Survey%20on%20gambling%20health%20and%20wellbeing%20in%20the%20ACT.pdf archived at http://www.webcitation.org/6k8kVQRKs
NSW	1995	Dickerson M, Allcock C, Blaszczynski A, Williams J, Maddern R. An examination of the socio-economic effects of gambling on individuals, families and the community, including research into the costs of problem gambling in New South Wales [Internet]. Sydney: Australian Institute for Gambling Research, University of Western Sydney; 1996. Available from: https://www.liquorandgaming.justice.nsw.gov.au/Documents/gaming-and-wagering/problems-with-gambling/research/2.%20Study%202%20%E2%80%93%20An%20Examination%20of%20the%20Socio-economic%20Effects%20of%20Gambling%20on%20Individuals,%20Families%20and%20the%20Community%20-%20part%201.pdf archived at http://www.webcitation.org/6k8kzUWQP
NSW	1997	Dickerson M, Allcock C, Blaszczynski A, Maddern R, Nicholls B, Williams J. An examination of the socio-economic effects of gambling on individuals, families and the community, including research into the costs of problem gambling (study 2 update) [Internet]. Sydney: Australian Institute for Gambling Research, University of Western Sydney; 1998. Available from: https://www.liquorandgaming.justice.nsw.gov.au/Documents/gaming-and-wagering/problems-with-gambling/research/1.%20An%20Examination%20of%20the%20Socio-Economic%20Effects%20of%20Gambling%20-%20part%201.pdf and https://www.liquorandgaming.justice.nsw.gov.au/Documents/gaming-and-wagering/problems-with-gambling/research/1.%20An%20Examination%20of%20the%20Socio-Economic%20Effects%20of%20Gambling%20-%20part%202.pdf archived at http://www.webcitation.org/6k8lSm0dN and http://www.webcitation.org/6k8lSwkR0

NSW	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E
NSW	2006	ACNielsen. Prevalence of gambling and problem gambling in NSW: A community survey 2006 [Internet]. ACNielsen; 2007. Available from: https://www.liquorandgaming.justice.nsw.gov.au/Documents/gaming-and-wagering/problems-with-gambling/research/8.%20Prevalence%20of%20Gambling%20and%20Problem%20Gambling%20in%20NSW%20%E2%80%93%20A%20Community%20Survey%202006.pdf archived at http://www.webcitation.org/6k8m1TlZC
NSW	2008-09	NSW Department of Health. Gambling Module, NSW Population Health Survey, 2008-2009 [Internet]. Sydney: NSW Department of Health; 2009. Available from: https://www.liquorandgaming.justice.nsw.gov.au/Documents/gaming-and-wagering/problems-with-gambling/research/1.%20Gambling%20Module%20NSW%20Population%20Health%20Survey%202008-2009%20-%20February%202010.pdf archived at http://www.webcitation.org/6k8mCbGGY
NSW	2011	Sproston K, Hing N, Palankay C. Prevalence of gambling and problem gambling in New South Wales [Internet]. Ogilvy Illumination; 2012. Available from: https://www.liquorandgaming.justice.nsw.gov.au/Documents/gaming-and-wagering/problems-with-gambling/research/7.%20Prevalence%20of%20gambling%20and%20problem%20gambling%20in%20NSW%20-%20October%202012.pdf archived at http://www.webcitation.org/6k8mw0C4r
NT	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E
NT	2005	Young M, Abu-Duhou I, Barnes T, Creed E, Morris M, Stevens M, et al. Northern Territory gambling prevalence survey 2005 [Internet]. Darwin: School for Social and Policy Research, Charles Darwin University; 2006. Available from: https://business.nt.gov.au/__data/assets/pdf_file/0010/275734/nt-gambling-prevalence-survey-2006.pdf archived at http://www.webcitation.org/6k8nEAPvy
QLD	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E
QLD	2001	Department of Corrective Services, Queensland Government. Problem gambling prevalence survey 2002 [Internet]. Brisbane: Queensland Government; 2002. Available from: https://publications.qld.gov.au/storage/f/2014-06-20T02%3A36%3A55.113Z/problem-gambling-prevalence-survey-2002.pdf archived at http://www.webcitation.org/6k8ngnmfp
QLD	2003-04	Queensland Government. Queensland household gambling survey 2003-04 [Internet]. Brisbane: Queensland Government; 2007. Available from: https://publications.qld.gov.au/storage/f/2014-06-20T02%3A43%3A47.596Z/queensland-household-gambling-survey-2003-04.pdf archived at http://www.webcitation.org/6k9ukUDwQ

QLD	2006-07	Queensland Government. Queensland household gambling survey 2006-07 [Internet]. Brisbane: Queensland Government; 2008. Available from: https://publications.qld.gov.au/storage/f/2014-06-20T02%3A41%3A11.146Z/queensland-household-gambling-survey-2006-07.pdf archived at http://www.webcitation.org/6k9uw4x8N
QLD	2008-09	Queensland Government. Queensland household gambling survey 2008-09 [Internet]. Brisbane: Queensland Government; 2010. Available from: https://publications.qld.gov.au/storage/f/2014-06-20T02%3A40%3A21.071Z/queensland-household-gambling-survey-2008-09.pdf archived at http://www.webcitation.org/6k9v6PRqm
QLD	2011-12	Queensland Government. Queensland household gambling survey 2011-12 [Internet]. Brisbane: Queensland Government; 2012. Available from: https://publications.qld.gov.au/storage/f/2014-06-20T02%3A38%3A40.297Z/queensland-household-gambling-survey-2011-12.pdf archived at http://www.webcitation.org/6k9vFTo4d
SA	1996	Delfabbro P, Winefield AH. Community gambling patterns and the prevalence of gambling-related problems in South Australia (with particular reference to gaming machines). Adelaide: Department of Psychology, University of Adelaide; 1996.
SA	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E
SA	2001	Taylor A, Dal Grande E, Gill T, Delfabbro P, Glenn V, Goulding S, et al. Gambling patterns of South Australians and associated health indicators [Internet]. Adelaide: South Australian Department of Human Services; 2001. Available from: https://health.adelaide.edu.au/pros/docs/reports/general/gambling_patterns_sa.pdf archived at http://www.webcitation.org/6k9wUxJXT
SA	2005	Taylor A, Gill T, Head L, Pugh L, Starr G, Rogers N, et al. Gambling prevalence in South Australia [Internet]. Adelaide: Department for Families and Communities, Government of South Australia; 2006. Available from: http://www.treasury.sa.gov.au/__data/assets/pdf_file/0015/1545/prevelance-study-report.pdf archived at http://www.webcitation.org/6k9wpKs9y
SA	2012	The Social Research Centre. Gambling prevalence in South Australia (2012) [Internet]. North Melbourne: The Social Research Centre; 2013. Available from: http://www.problemgambling.sa.gov.au/professionals/news_and_events/news-items/release-of-the-2012-gambling-prevalence-study-in-south-australia?a=13625 archived at http://www.webcitation.org/6k9x2DJAk
TAS	1994	Dickerson M, Walker M, Baron E. A baseline study of the extent and impact of gambling in Tasmania with particular reference to problem gambling. Sydney: Australian Institute for Gambling Research, University of Western Sydney; 1994.
TAS	1996	Dickerson M, Maddern R. The extent and impact of gambling in Tasmania with particular reference to problem gambling: A follow up to the baseline study conducted 1994. Sydney: Australian Institute for Gambling Research, University of Western Sydney; 1997.

TAS	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E
TAS	2000	Roy Morgan Research. The third study into the extent and impact of gambling in Tasmania with particular reference to problem gambling : follow up to the baseline studies conducted in 1994 and 1996. Melbourne: Roy Morgan Research; 2001.
TAS	2005	Roy Morgan Research. The fourth study into the extent and impact of gambling in Tasmania with particular reference to problem gambling: Follow up to the studies conducted in 1994, 1996 and 2000. Melbourne: Roy Morgan Research; 2006.
TAS	2007	The SA Centre for Economic Studies. Social and economic impact study into gambling in Tasmania: Volume 2: The Prevalence Study [Internet]. Adelaide: University of Adelaide; 2008. Available from: http://www.treasury.tas.gov.au/domino/dtf/dtf.nsf/LookupFiles/Soc-Economic-Impact-Study-Vol2.pdf/$file/Soc-Economic-Impact-Study-Vol2.pdf archived at http://www.webcitation.org/6kA1g9Y8b
TAS	2011	The Allen Consulting Group, Problem Gambling Research and Treatment Centre, The Social Research Centre. Social and Economic Impact Study of Gambling in Tasmania: Volume 2: Gambling survey. Melbourne: Allen Consulting Group; 2011 p. http://www.treasury.tas.gov.au/domino/dtf/dtf.nsf/LookupFiles/Volume2secondgamblingSEIS.PDF/$file/Volume2secondgamblingSEIS.PDF archived at http://www.webcitation.org/6kA1Wrptl
TAS	2013	The Allen Consulting Group, The Social Research Centre, The Problem Gambling Research and Treatment Centre. Third social and economic impact study of gambling in Tasmania: Volume 2: 2013 Tasmanian gambling prevalence survey [Internet]. Melbourne: Allen Consulting Group; 2014. Available from: http://www.treasury.tas.gov.au/domino/dtf/dtf.nsf/LookupFiles/20150109SEISVolume2FINALREVISEDCHANGES.PDF/$file/20150109SEISVolume2FINALREVISEDCHANGES.PDF archived at http://www.webcitation.org/6kA1rNGoe
VIC	1997	Maddern C, Horman S, Dickerson M. Fifth community gambling patterns survey combined with second positive and negative perceptions of gambling survey. Melbourne: Victorian Casino and Gaming Authority; 1997.
VIC	1998	Roy Morgan Research. Sixth survey of community gambling patterns and perceptions [Internet]. Melbourne: Roy Morgan Research; 1999. Available from: http://assets.justice.vic.gov.au/vcglr/resources/9412cce8-63fe-4f5e-82e0-fe5c0951b62a/6thsurveycommunitygamblingpattersnandperceptions.pdf archived at http://www.webcitation.org/6kA2LeY36
VIC	1999a	Roy Morgan Research. Seventh survey of community gambling patterns and perceptions [Internet]. Melbourne: Roy Morgan Research; 2000. Available from: http://assets.justice.vic.gov.au/vcglr/resources/4ea71095-8ecc-44e4-a096-5477dfaf4f06/7thsurveycommunitygamblingpatternsandperceptions.pdf archived at http://www.webcitation.org/6kA2dQikr
VIC	1999b	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E

VIC	2003	McMillen J, Marshall D, Ahmed E, Wenzel M. 2003 Victorian Longitudinal Community Attitudes Survey [Internet]. Canberra: The Centre for Gambling Research, The Australian National University; 2004. Available from: http://hdl.handle.net/1885/45189 archived at http://www.webcitation.org/6kA31QkOI
VIC	2007	Thomas SA, Jackson AC. Risk and protective factors, depression and comorbidities in problem gambling. Melbourne: beyondblue; 2008.
VIC	2008	Hare S. A study of gambling in Victoria: Problem gambling from a public health perspective [Internet]. Melbourne: Department of Justice, Government of Victoria; 2009. Available from: https://www.responsiblegambling.vic.gov.au/__data/assets/pdf_file/0013/4027/A-study-of-gambling-in-Victoria-PRINT-Sept-10.pdf archived at http://www.webcitation.org/6kA3s64IZ
VIC	2014	Hare S. Study of gambling and health in Victoria: Findings from the Victorian prevalence study 2014 [Internet]. Melbourne: Victorian Responsible Gambling Foundation; 2015. Available from: https://www.responsiblegambling.vic.gov.au/__data/assets/pdf_file/0018/25551/Study_of_gambling_and_health_in_Victoria.pdf archived at http://www.webcitation.org/6kA43TuMs
WA	1994	Dickerson M, O'Connor J, Baron E. An assessment of the extent and degree of gambling related problems in the population of Western Australia. Sydney: Australian Institute for Gambling Research, University of Western Sydney; 1994.
WA	1999	Productivity Commission. Australia's Gambling Industries. Report no.: 10.Canberra: Productivity Commission; 1999. Available from: http://www.pc.gov.au/inquiries/completed/gambling/report/gambling1.pdf archived at http://www.webcitation.org/6k8jP5l5E

Appendix F: Meta-analyses of methodological variations for Chapter 5

Introduction

Many national and subnational jurisdictions undertake problem gambling prevalence studies to estimate the number of problem gamblers in the adult population. Indeed, over 200 problem gambling prevalence studies were conducted between 1975 and 2012 (Williams, Volberg, and Stevens, 2012), a number that continues to grow as some jurisdictions conduct their first ever problem gambling prevalence studies and others repeat the exercise in order to monitor change over time.

As Markham and Young (2016) remark, these studies have at least three objectives:

1. to assess the burden of disease in a population and to assess the need for health services;
2. to compare the prevalence of disease in different populations; and
3. to examine trends in disease prevalence or severity over time.

In practice, objectives 2 and 3 generally involve the comparison of the results of different problem gambling prevalence studies, whether this is done to compare prevalence between jurisdictions or to assess change within a single jurisdiction over time. However, as many authors have noted, such comparisons are problematic due to methodological variations between problem gambling prevalence studies (e.g. Williams & Volberg, 2009, 2010; Jackson *et al.*, 2010; Sassen, Kraus, and Bühringer, 2011; Stone et al. 2014).

This is especially problematic as the impact of methodological variations on prevalence estimates is likely to be of a greater magnitude than actual variations in underlying population prevalence. For example, a relative increase of 10% in the true prevalence of problem gambling is likely to be large enough to be policy relevant. However, as Williams and Volberg (2009) show, the impact of methodological variations on prevalence estimates may be on the order of 100% in isolation, and more than 400% when combined. Therefore, the correction for any methodological variations is important when comparing the results of prevalence studies.

Objectives

This paper undertakes meta-analyses of previous studies of the methodological variations among prevalence studies to estimate the magnitude of the corrections that must be made prior to comparing problem gambling prevalence studies. Specifically, this paper seeks to estimate the average impact of the following methodological variations of prevalence estimates:

- choice of problem gambling screen: whether problem gambling was assessed using the South Oaks Gambling Screen (SOGS) or the Problem Gambling Severity Index (PGSI)
- administration mode: whether the questionnaire was administered by telephone or face-to-face
- frequency threshold: if the problem gambling screen was administered to everyone, only to those who gambled in the last twelve months, only those who gambled in the last month, or only those who gambled in the last week.

For each of these three methodological variations, this study aims to:

1. Identify previous research on these topics
2. Summarise previous estimates of these methodological variations on problem gambling prevalence estimates
3. Describe the heterogeneity among previous estimates

Methods

Random effects meta-analytical approach was taken to answer these research questions. Three separate analyses were undertaken in order to identify and summarise previous research on these three methodological variations in problem gambling prevalence studies.

Problem gambling screen search strategy

Studies were eligible for inclusion in this meta-analysis if they were primary studies that compared and reported problem gambling prevalence estimates using both the PGSI and the SOGS. This comparison could either take place either by administering both screens to the same respondents, or by administering different screens to two random subsamples of the same sample. Studies were excluded if they applied to different time periods (i.e. if they administered the PGSI for the period of the last twelve months and SOGS over the lifetime).

Studies were identified by searching Scopus, the Web of Knowledge and by examining Williams, Volberg and Stevens' (2012) listing of problem gambling prevalence studies.

The following search terms were used in online databases:

Web of Science

```
TOPIC: (sogs AND pgsi)
Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-
SSH, ESCI, CCR-EXPANDED, IC.
```

Scopus

```
TITLE-ABS-KEY ( sogs  AND  pgsi )
```

Using the listing of problem gambling prevalence studies from Williams, Volberg and Stevens (2012), studies were identified that administered both the SOGS and PGSI.

Studies were screened by FM to ensure that the eligibility criteria were met, and data extracted into a spreadsheet. Where prevalence of problem gambling was reported as a percentage, these estimates were converted to a count of cases. The following data items were extracted for each study:

1. Whether both the SOGS and the PGSI were administered to respondents, or whether different subgroups of respondents were administered different screens
2. The count of respondents endorsing 5 more SOGS items
3. The denominator for this count
4. The count of respondents scoring 8 more on the PGSI
5. The denominator for this count
6. The count of respondents endorsing 3-4 SOGS items
7. The denominator for this count
8. The count of respondents scoring 3-7 on the PGSI
9. The denominator for this count

Administration mode search strategy

Studies were eligible for inclusion in this meta-analysis if they were primary studies which compared and reported problem gambling prevalence estimates from surveys administered by telephone and face-to-face. Only experiments which administered a common survey instrument to a single sample frame, using random selection to assign respondents to an administration mode were eligible for inclusion. Studies which compared telephone surveys with postal surveys were ineligible.

Studies were identified by searching Scopus and the Web of Knowledge and by examining Williams, Volberg and Stevens' (2012) listing of problem gambling prevalence studies.

The following search terms were used in online databases:

Web of Science

```
TOPIC:
(gambling OR "problem gambling" OR "pathological gambling" OR pgsi OR
cpgi OR sogs) AND
TOPIC: (( "administration mode" OR "administration format" OR "face-
to-face"))
Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-
SSH, ESCI, CCR-EXPANDED, IC.
```

Scopus

```
( TITLE-ABS-KEY ( gambling  OR  "problem gambling"  OR  "pathological
gambling"  OR  pgsi  OR  cpgi  OR  sogs )  AND  TITLE-ABS-KEY (
"administration mode"  OR  "administration format"  OR  "face-to-face"
) )
```

The full text of relevant studies was read for further references.

Studies were screened by FM to ensure that the eligibility criteria were met, and data extracted into a spreadsheet. Where prevalence of problem gambling was reported as a percentage, these estimates were converted to a count of cases. If results were reported with population weights applied and without population weights applied, both data values were extracted. The following data items were extracted for each study:

1. The count of respondents classified as problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) in the telephone survey
2. The denominator for this count
3. The count of respondents classified as problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) in the face-to-face survey
4. The denominator for this count

5. The count of respondents classified as moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4) in the telephone survey
6. The denominator for this count
7. The count of respondents classified as moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4) in the face-to-face survey
8. The denominator for this count

Frequency threshold search strategy

Studies were eligible for inclusion in this meta-analysis if they were primary studies that compared and reported problem gambling prevalence estimates made on the basis of different frequency thresholds for administering the problem gambling screen. This comparison could either take place either by administering both screens to the same respondents, or by administering different screens to two random subsamples of the same sample.

Studies were identified by searching Scopus, the Web of Knowledge and by examining Williams, Volberg and Stevens' (2012) listing of problem gambling prevalence studies. All studies in Williams, Volberg and Stevens' (2012) list were eligible if a) they were not the first problem gambling prevalence study in that jurisdiction, b) the previous study in that jurisdiction only administered the problem gambling screen to those respondents who met a gambling frequency threshold, and c) that study used no frequency threshold, or a lower frequency threshold. The logic behind these criteria is that in eligibility studies, a valid comparison of problem gambling prevalence with a past estimate in the same jurisdiction requires re-estimation using a different frequency threshold.

The following search terms were used in online databases:

Web of Science

```
TOPIC: (("frequency threshold" OR "sub-sample" OR "exclusion criteria"
OR "regular gamblers" OR "non-regular gamblers" OR "frequent gamblers"
OR "infrequent gamblers")) AND TOPIC: ((gambling OR "problem gambling"
OR "pathological gambling" OR pgsi OR cpgi OR sogs))
Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-
SSH, ESCI, CCR-EXPANDED, IC.
```

Scopus

```
( TITLE-ABS-KEY ( "frequency threshold" OR "sub-sample" OR
"exclusion criteria" OR "regular gamblers" OR "non-regular
gamblers" OR "frequent gamblers" OR "infrequent gamblers" ) AND
TITLE-ABS-KEY ( gambling OR "problem gambling" OR "pathological
gambling" OR pgsi OR cpgi OR sogs ) )
```

Using the exclusion criteria described above and the listing of problem gambling prevalence studies from Williams, Volberg and Stevens' (2012), studies were identified that reported problem gambling prevalence estimates based on two different frequency thresholds.

Studies were screened by FM to ensure that the eligibility criteria were met, and data extracted into a spreadsheet. Where prevalence of problem gambling was reported as a percentage, these estimates were converted to a count of cases. Where studies included other exclusion criteria in addition to a frequency threshold (e.g. they ignored lottery gambling when selecting a subsample) these other criteria were ignored. Studies that administered the problem gambling screen only to those who had gambled at least once in the past year or in their lifetime were grouped with studies that administered the screen to all respondents. The following data items were extracted for each study:

1. Whether the problem gambling screen was the SOGS or the PGSI
2. The count of problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) when the screen was administered only to those who gambled weekly or more frequently
3. The denominator for this count
4. The count of moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4) when the screen was administered only to those who gambled weekly or more frequently
5. The denominator for this count
6. The count of problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) when the screen was administered only to those who gambled fortnightly or more frequently
7. The denominator for this count
8. The count of moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4) when the screen was administered only to those who gambled fortnightly or more frequently
9. The denominator for this count
10. The count of problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) when the screen was administered only to those who gambled monthly or more frequently
11. The denominator for this count

12. The count of moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4) when the screen was administered only to those who gambled monthly or more frequently
13. The denominator for this count
14. The count of problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) when the screen was administered to everyone or those who had ever gambled or those who had gambled in the past year
15. The denominator for this count
16. The count of moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4) when the screen was administered to everyone or those who had ever gambled or those who had gambled in the past year
17. The denominator for this count

Statistical analysis

The principal summary measures are the ratios of prevalence estimates in the contrasting conditions. These are calculated as a risk ratio would be calculated in a conventional meta-analysis.

As it was anticipated that prevalence ratios would exhibit substantial heterogeneity, it was decided to use a random effects meta-analysis model. Each meta-analysis was repeated twice, once for problem gamblers (PGSI ≥ 8 or SOGS ≥ 5) and once for moderate risk problem gamblers (PGSI 3-7 or SOGS 3-4). As well as estimating mean prevalence ratios and their 95% confidence intervals, 95% prediction intervals were calculated in order to estimate the range of values expected in future studies of this type. τ and I^2, measures of the heterogeneity of studies were also calculated and reported in each meta-analysis. No assessment of the risk of bias was made for individual studies. No review protocol was registered for this study.

Results

Problem gambling screen

A total of 237 studies were screened, yielding 12 studies that met the eligibility criteria (see Figure E.1). Characteristics of individual studies included are listed in Table E.1.

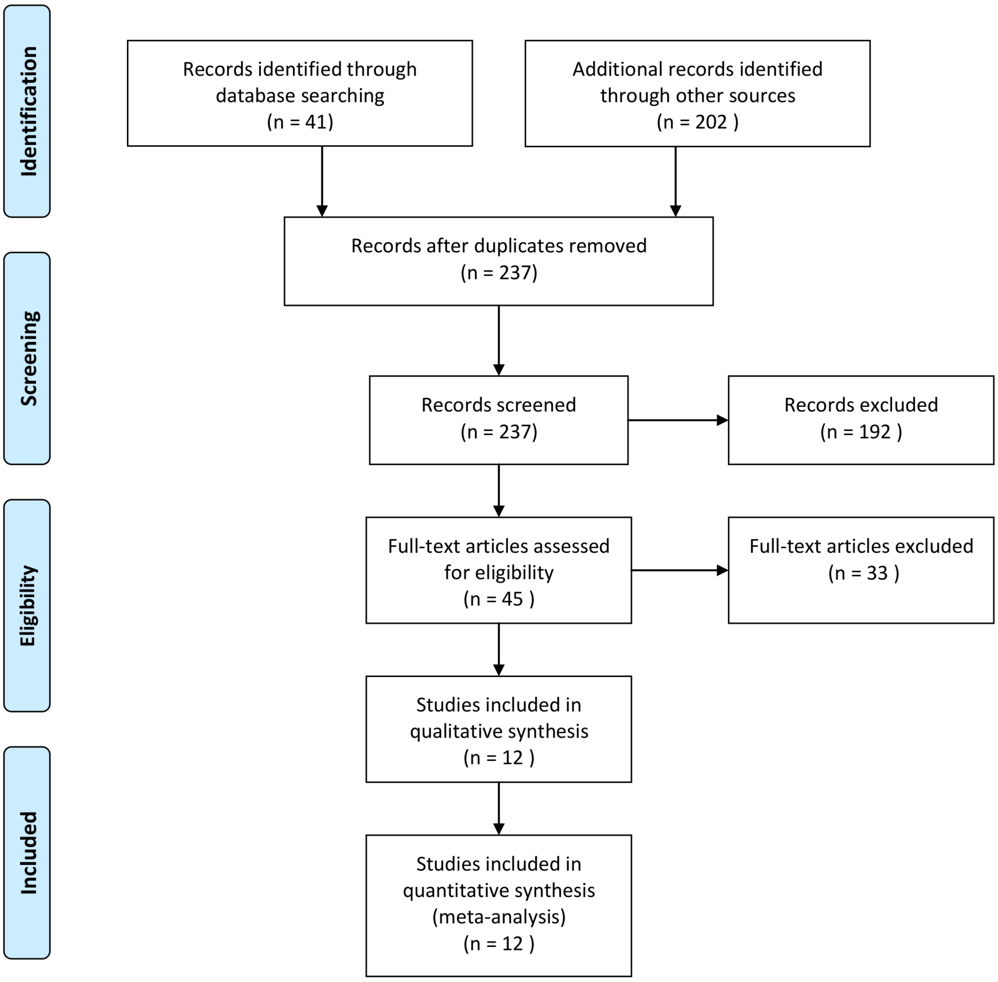

Figure F.1: Literature search flow diagram for problem gambling screen meta-analysis

Table F.1: Studies of problem gambling screen effects from which data were extracted

Study	Same respondents	SOGS 5+	SOGS *n*	PGSI 8+	PGSI *n*	SOGS 3-4	PGSI 3-7
NT 2005	TRUE	54	369	38	369	50	68
VIC 2003	FALSE	26	143	22	141	22	21
BPDS 2010	TRUE	32	2193	13	2193	25	77
IODS 2010	TRUE	664	5078	467	5079	730	1247
Ferris and Wynne 2001	TRUE	41	3120	28	3120	41	74
British Columbia 2002	TRUE	23	2134	10	2500	58	105
Manitoba 2001	TRUE	72	3119	34	3119	NA	72
Italy 2008	TRUE	69	1987	73	1987	64	233
Sweden 2009	TRUE	120	15000	45	15000	180	285
Hungary 2013	TRUE	78	586	37	586	75	68
Quebec 2002	FALSE	41	4603	30	4225	41	42
Turner et al. 2009	TRUE	33	254	24	254	12	40

As Figure E.2 shows, the mean ratio of problem gambling prevalence estimates made with SOGS ≥ 5 and PGSI ≥ 8 was 1.6 (95% C.I. 1.3, 2.0), meaning that studies conducted using SOGS identified an average of 1.6 times more problem gamblers. The 95% prediction interval around this estimate was 0.9 - 2.8. A τ value of 0.27 (95% C.I. 0.12, 0.54) was recorded, and an I^2 of 69% (95% C.I. 32%, 90%), a moderate degree of heterogeneity.

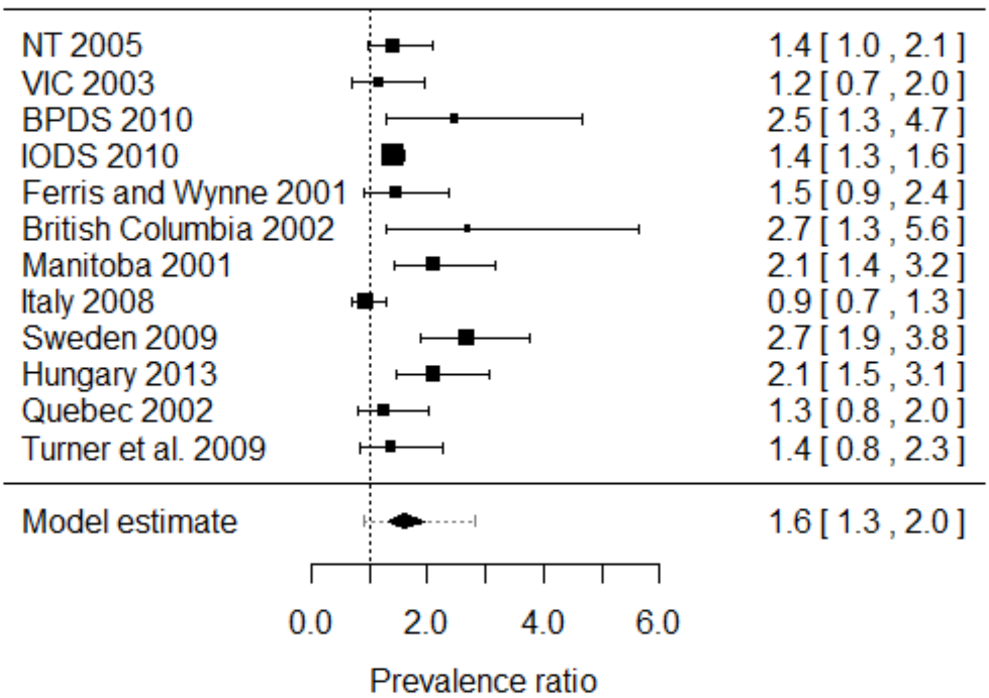

Figure F.2: SOGS 5+ to PGSI 8+ prevalence ratios and model estimate. Dotted line indicates 95% prediction interval.

The prevalence ratio is quite different for moderate risk problem gambling (Figure E.3). The mean ratio of the prevalence of SOGS 3-4 and PGSI 3-7 was 0.6 (95% C.I. 0.4, 0.8), meaning that studies conducted using SOGS identified an average of 0.6 times fewer people. The 95% prediction interval around this estimate was 0.2 - 1.4. A τ value of 0.42 (95% C.I. 0.27, 0.82) was recorded, and an I^2 of 91% (95% C.I. 80%, 97%), a very high degree of heterogeneity.

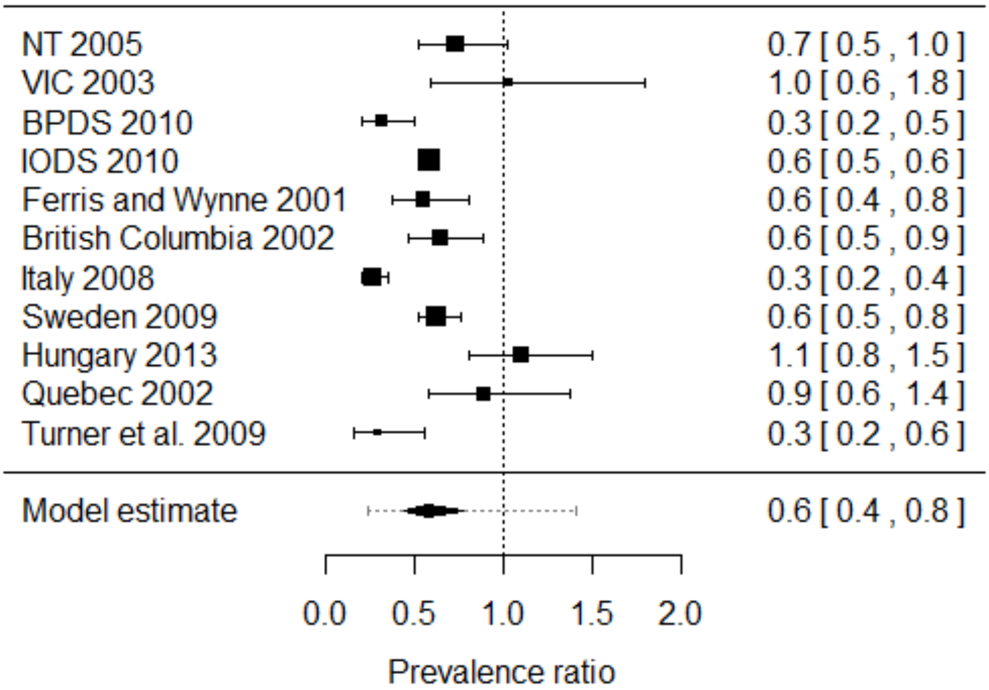

Figure F.3: SOGS 3-4 to PGSI 3-7 prevalence ratios and model estimate. Dotted line indicates 95% prediction interval.

Administration mode

Although 81 studies were screened, only a single study that met the eligibility criteria was identified (see Figure E.4). That study published two estimates of the effect of administration, one with population weights applied and another unweighted estimate. The results of these two estimates are listed in Table E.2.

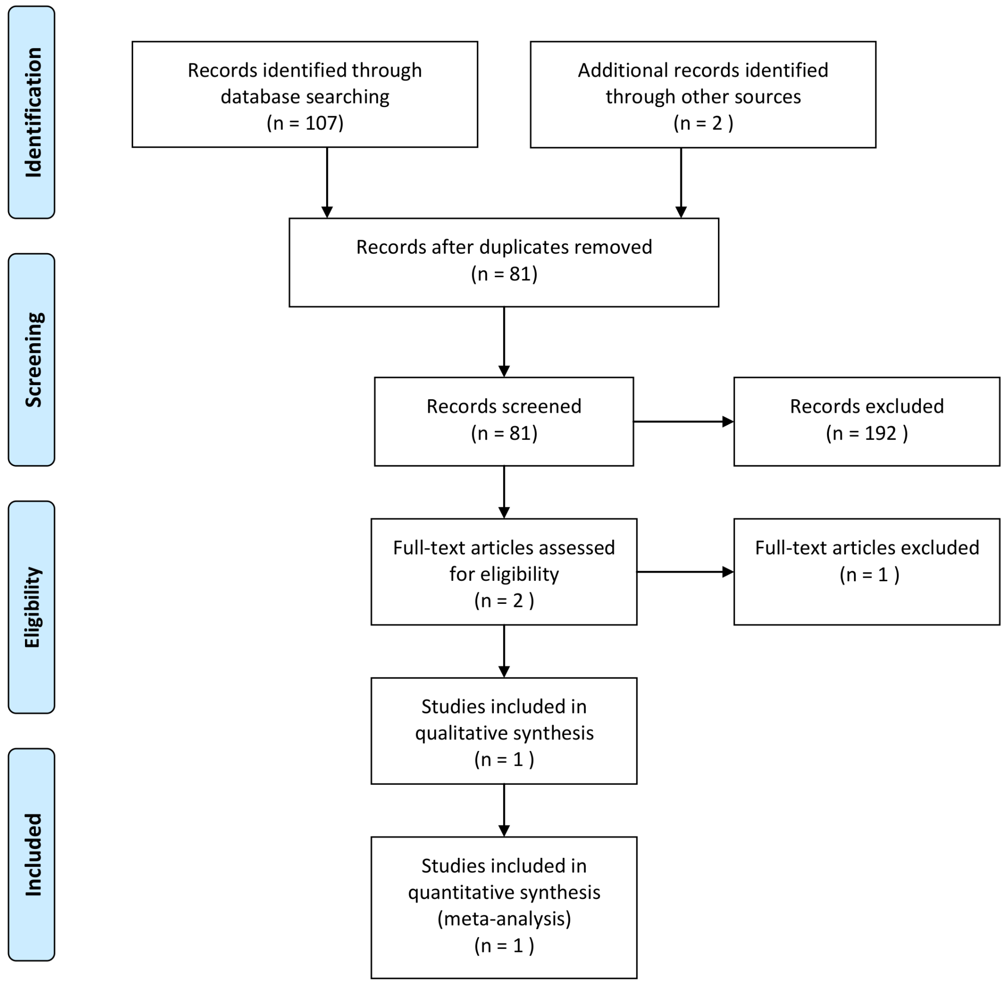

Figure F.4: Literature search flow diagram for administration mode meta-analysis

Table F.2: Estimates from which data were extracted

	Problem gambling				Moderate risk problem gambling			
Study	Telephone	Telephone *n*	Face to face	Face to face *n*	Telephone	Telephone *n*	Face to face	Face to face *n*
W & V 2009 unweighted	6	1513	7	1515	22	1513	54	1515
W & V 2009 weighted	6	1518	6	1528	24	1518	42	1528

As Figure E.5 shows, the mean ratio of problem gambling prevalence estimates administered by doorknock compared to telephone was 1.1 (95% C.I. 0.5, 2.4), meaning that studies administered face-to-face identified an average of 1.1 times more problem gamblers. As only two reports of a single study were identified, heterogeneity was unable to be estimated.

Results were much more extreme for moderate risk problem gamblers (Figure E.6). The mean ratio of moderate risk prevalence estimates administered by doorknock compared to telephone was 2.1 (95% C.I. 1.5, 2.9), meaning that studies administered face-to-face identified an average of 2.1 times more moderate risk gamblers.

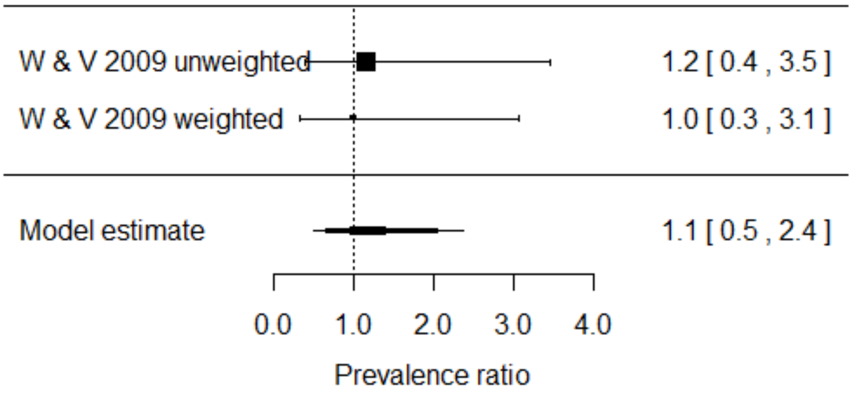

Figure F.5: Doorknock administered to telephone administered prevalence ratios and model estimate for PGSI >= 8.

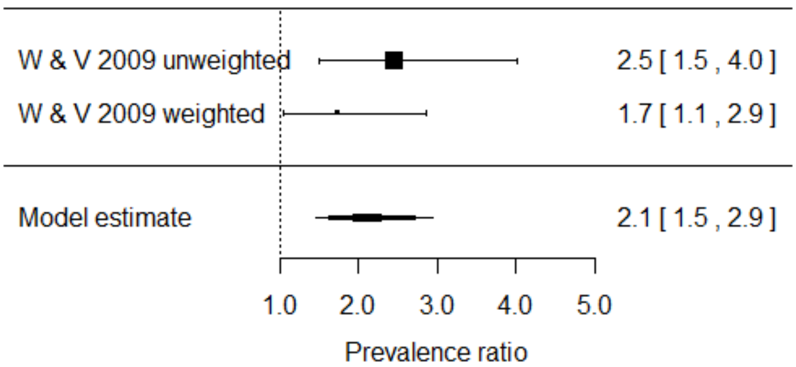

Figure F.6: Doorknock administered to telephone administered prevalence ratios and model estimate for PGSI 3-7.

Frequency threshold

A total of 322 studies were screened, yielding 5 studies that met the eligibility criteria (see Figure E.7). Characteristics of individual studies included are listed in Table E.3.

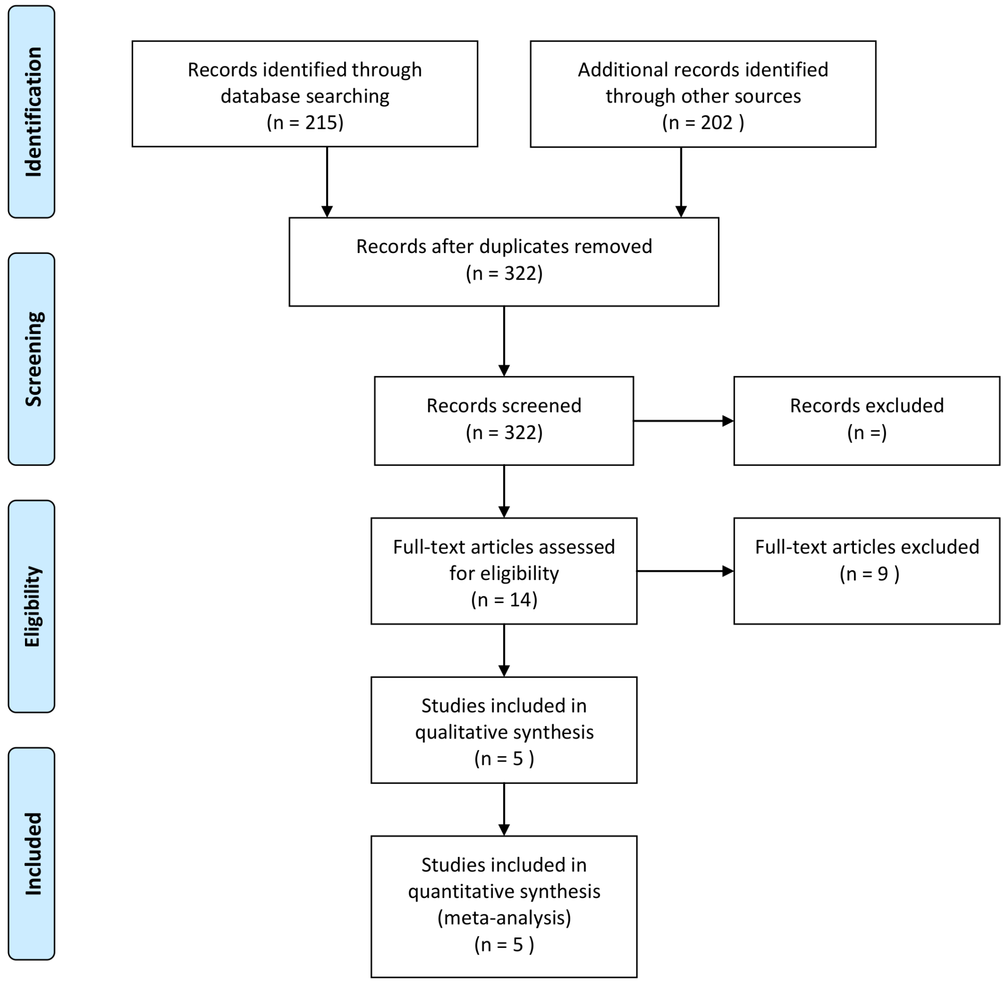

Figure F.7: Literature search flow diagram for frequency threshold meta-analysis

Table F.3: Estimates from which data were extracted

Study	Screen	Weekly threshold			Fortnightly threshold			Monthly threshold			Annual or no threshold		
		Problem gamblers	Moderate risk problem gamblers	*n*	Problem gamblers	Moderate risk problem gamblers	*n*	Problem gamblers	Moderate risk problem gamblers	*n*	Problem gamblers	Moderate risk problem gamblers	*n*
NSW 2011	PGSI	40	150	10000							80	290	10000
SA 2012	PGSI				49	143	9246				55	231	9246
TAS 1996	SOGS	27	42	1211							35	69	1211
TAS 2011	PGSI	17	30	4303							30	77	4303
Jackson, Wynne et al 2009 (VIC 2007)	PGSI	18	39	1488							28	56	1488
Stone et al. - VIC 2008	PGSI	81	195	15000	93	233	15000	101	300	15000	105	350	15000
Stone et al. - Swelogs 2008	PGSI	36	185	15000	47	213	15000	47	246	15000	48	281	15000

As Figure E.8 shows, the mean ratio of problem gambling prevalence estimates made with a weekly frequency threshold compared to an annual frequency threshold was 0.7 (95% C.I. 0.6, 0.8), meaning that studies conducted using a weekly frequency threshold identified an average of 0.7 times fewer problem gamblers. The 95% prediction interval around this estimate was 0.5 - 0.8 A τ value of 0.06 (95% C.I. 0.00, 0.39) was recorded, and an I^2 of 7% (95% C.I. 0%, 76%), a very small degree of heterogeneity.

Less of an effect can be detected for fortnightly frequency thresholds, perhaps due to the smaller number of studies that have analysed them. Figure E.9 shows that the mean ratio of problem gambling prevalence estimates made with a fortnightly frequency threshold compared to an annual frequency threshold was 0.9 (95% C.I. 0.7, 1.1), meaning that studies conducted using a fortnightly frequency threshold identified an average of 0.9 times fewer problem gamblers. This is not significantly different to no effect. Heterogeneity could not be estimated for fortnightly thresholds.

A similar result was found for monthly frequency thresholds. Figure E.10 shows just two studies of these thresholds, with a mean prevalence ratio of 1.0 (95% C.I. 0.8, 1.2). This is not significantly different to no effect. Heterogeneity could not be estimated for fortnightly thresholds.

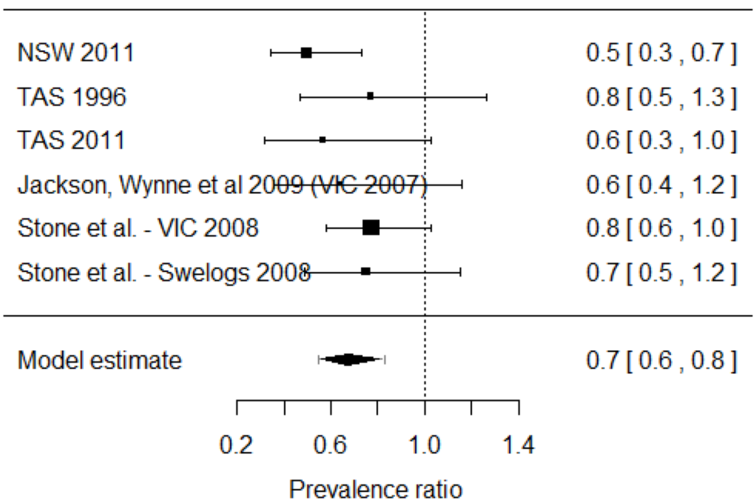

Figure F.8: Ratio of problem gambling prevalence estimates for weekly frequency thresholds compared to annual frequency thresholds. Dotted line indicates 95% prediction interval.

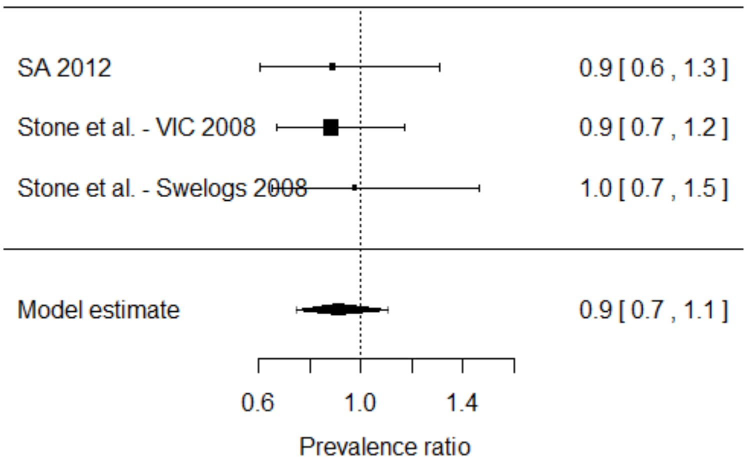

Figure F.9: Ratio of problem gambling prevalence estimates for fortnightly frequency thresholds compared to annual frequency thresholds. Dotted line indicates 95% prediction interval.

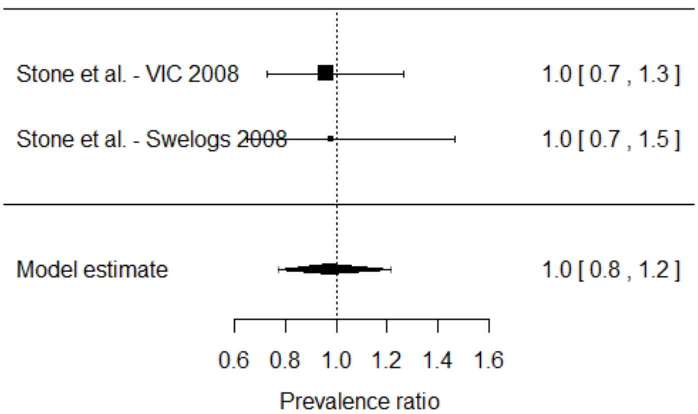

Figure F.10: Ratio of problem gambling prevalence estimates for monthly frequency thresholds compared to annual frequency thresholds. Dotted line indicates 95% prediction interval.

Frequency thresholds appear to have a stronger effect on moderate risk problem gambling (see Figures E.11-E.13). For moderate risk problem gamblers (PGSI 3-7 and SOGS 3-4), the effect of weekly, fortnightly and monthly frequency thresholds were estimated in terms of prevalence ratios as 0.6 (95% C.I. 0.5, 0.6), 0.7 (95% C.I. 0.6, 0.8) and 0.9 (95% C.I. 0.8, 1.0), respectively.

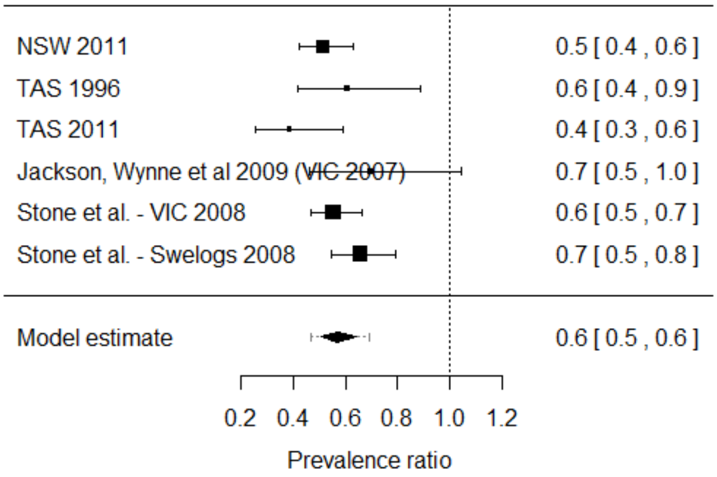

Figure F.11: Ratio of moderate risk problem gambling prevalence estimates for weekly frequency thresholds compared to annual frequency thresholds. Dotted line indicates 95% prediction interval.

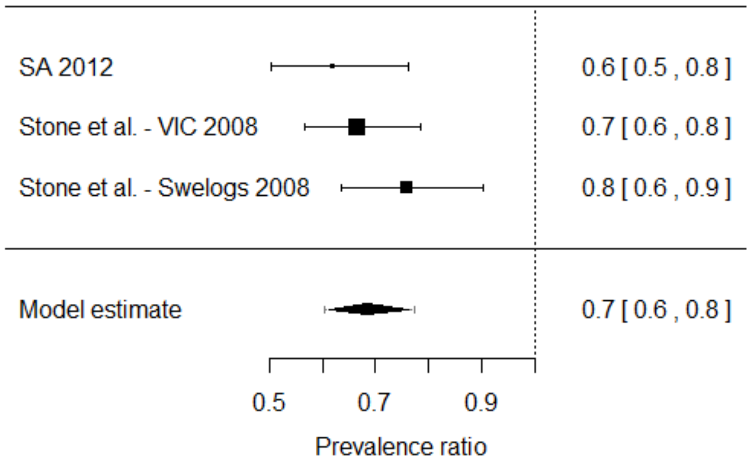

Figure F.12: Ratio of moderate risk problem gambling prevalence estimates for fortnightly frequency thresholds compared to annual frequency thresholds. Dotted line indicates 95% prediction interval.

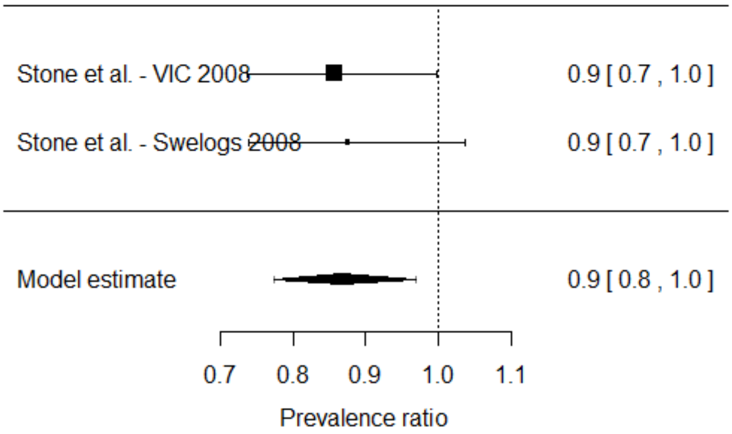

Figure F.13: Ratio of moderate risk problem gambling prevalence estimates for monthly frequency thresholds compared to annual frequency thresholds. Dotted line indicates 95% prediction interval.

Discussion and conclusions

In the case of problem gambling, there is substantial evidence that screen choice has a significant impact on prevalence estimates. More precisely, estimates using SOGS ≥ 5 identify 1.6 (95% C.I. 1.3, 2.0) times more problem gamblers than studies using PGSI ≥ 8. Only a single study of the effect of administration mode on problem gambling prevalence estimates could be identified, and it found no evidence of a significant effect. Studies which administered a problem gambling screen only to weekly gamblers identified an average of 0.7 (95% C.I. 0.6, 0.8) times fewer problem gamblers. No significant effect on problem gambler prevalence was found for fortnightly or monthly thresholds, but this could be due to the small number of studies that investigated these thresholds.

For moderate risk problem gambling, defined as a SOGS score of 3-4 or a PGSI score of 3-7, results were quite different. Specifically, studies using SOGS 3-4 found 0.6 (95% C.I. 0.4, 0.8) times fewer moderate risk problem gamblers, compared with those using PGSI 3-7. While administration mode appeared to matter little for problem gambling, for moderate risk problem gambling a significant effect was found. Studies administered face-to-face identified 2.1 (95% C.I. 1.5, 2.9) times more moderate risk problem gamblers than those administered by telephone. Similarly, frequency thresholds had a greater effect on estimates of the prevalence of moderate risk problem gambling than problem gambling. Specifically, for moderate risk problem gamblers

(PGSI 3-7 and SOGS 3-4), the prevalence ratios associated with weekly, fortnightly and monthly frequency thresholds were 0.6 (95% C.I. 0.5, 0.6), 0.7 (95% C.I. 0.6, 0.8) and 0.9 (95% C.I. 0.8, 1.0), respectively. It seems likely that the reason why SOGS may be identifying fewer people at moderate risk of problem gambling is that it identified more people who were problem gamblers. In other words, individuals who may be classified as problem gamblers using the PGSI are iedentified as moderate risk using the SOGS.

Due to the small number of studies meeting the inclusion criteria, these results should be interpreted with caution. In particular, only a single study of the effect of administration mode on prevalence estimates was found, and only two studies investigated the impact of fortnightly and monthly frequency thresholds. Furthermore, the extreme heterogeneity was evident in estimates of the effect of screen choice on the prevalence of moderate risk problem gambling, meaning that estimates of these effects vary substantially from study to study.

Appendix G: JAGS models for Chapter 5

All JAGS model specification code is reproduced in listings G.1 – G.6 on the following pages.

Listing G.1: Random effects meta-analysis of the prevalence of problem gambling, without moderators

```
model
      {
         for(i in 1:N) {
            y[i] ~ dbin(p[i], s_size[i])
            logit(p[i]) <- intercept + theta[i]

            # Calculate the weights to help with estimating Q
        w[i] <- 1 / (1/y[i] + 1/(s_size[i] - y[i]))

            theta[i] ~ dnorm(0, prec)
            resid[i] <- y[i] - fitted_est[i]
            fitted_est[i] <- p[i] * s_size[i]
      }

      # Calculate Q from Borenstein et al, p. 114, eq. 16.5
      N_params <- 1

      # Calculate Q from eq 16.1 on p. 109, Borenstein et al.
      Q <- sum(w * ((logit(p) - intercept)^2))

      # Calculate I2
      # As defined on p. 1546, eq. 10 in http://doi.org/10.1002/sim.1186
      H.sq <- Q / (N - N_params)

      # As defined on Borenstein et al., p. 117, eq. 16.9
      I.sq <- min((Q - (N - N_params)) / max(Q, 10^-12), 1.0) * 100.0

      intercept ~ dnorm(0, 10^-6)

      tau~  dunif(0,10)  # Suggested by http://dx.doi.org/10.1002/sim.2112
      tau.sq <- tau*tau
      prec <- 1/(tau.sq)
```

```
      }


# Convenience settings for runjags in R:
#monitor# intercept, tau, tau.sq, Q, I.sq, H.sq
#modules# glm on
#response# y
#residual# resid
#fitted# fitted_est


####################################################################################################
####################################################################################################
#### Initial values
####################################################################################################
####################################################################################################

inits{
"intercept" <- -1.5
}

inits{
"intercept" <- -0.5
}

inits{
"intercept" <- 0.5
}

inits{
"intercept" <- 1.5
}
```

Listing G.2: Random effects meta-analysis of the prevalence of problem gambling, with moderators and informative priors

```
model
      {

         for(i in 1:N) {
            y[i] ~ dbin(p[i], s_size[i])
            logit(p[i]) <- intercept +
                  B_doorknock * doorknock[i] +
                  B_SOGS * SOGS[i] +
                  B_freq_thresh_m * freq_thresh_m[i] +
                  B_freq_thresh_f * freq_thresh_f[i] +
                  B_freq_thresh_w * freq_thresh_w[i] +
                  B_Years * Years_before_2016[i] +
                  B_Exp_EGM_pc_hdi * Exp_TG_EGM_pc_hdi[i] +
                  theta[i]

            theta[i] ~ dnorm(0, prec)

            # Inverse variance weights for I2
            w[i] <- 1 / (1/y[i] + 1/(s_size[i] - y[i]))

            # Repeat with no covariates for R2 calc
            y_full[i] ~ dbin(p_full[i], s_size[i])
            logit(p_full[i]) <- intercept_full + theta_full[i]
            theta_full[i] ~ dnorm(0, prec_full)

            # calculate residuals and fitted values to bring back into R
            resid[i] <- y[i] - fitted_est[i]
            fitted_est[i] <- p[i] * s_size[i]
      }

      # Priors
      FLATTEN_FACTOR <- 4
```

```
intercept ~ dnorm(0, 10^-6)
B_doorknock ~ dnorm(0.0761,(1/(0.3998*FLATTEN_FACTOR))^2)
B_SOGS ~ dnorm(0.4797, (1/(0.1000*FLATTEN_FACTOR))^2)
B_freq_thresh_m ~ dnorm(-0.0332, (1/(0.1150*FLATTEN_FACTOR))^2)
B_freq_thresh_f ~ dnorm( -0.0958, (1/(0.1002*FLATTEN_FACTOR))^2)
B_freq_thresh_w ~ dnorm(-0.3965, (1/(0.0915*FLATTEN_FACTOR))^2)
B_Years ~ dnorm(0, 10^-6)
B_Exp_EGM_pc_hdi ~ dnorm(0, 10^-6)

# Heterogeneity statistics
N_params <- 8

# Calculate tau and tau2
prec <- 1/(tau.sq)
tau.sq <- tau^2
tau ~ dunif(0,10)  # Suggested by http://dx.doi.org/10.1002/sim.2112

# Calculate I2
# As defined on p. 1546, eq. 9 and 10 http://doi.org/10.1002/sim.1186
H.sq <- (tau.sq + sigma.sq) / sigma.sq
sigma.sq <- (sum(w) * (N - 1)) / (sum(w)^2 - sum(w^2))
I.sq <- tau.sq / (tau.sq + sigma.sq)

# H.sq.v2 <- (((sum(w)  - (sum(w^2) / sum(w))) * tau.sq) / (N - 1)) + 1 # eq. 11 and 10 are equivalent
# I.sq.v2 <- (H.sq.v2 - 1) / H.sq.v2

# R2 calc
intercept_full ~ dnorm(0, 10^-6)
tau_full ~ dunif(0,10)
tau.sq_full <- tau_full^2
prec_full <- 1/(tau.sq_full)
R.sq <- (1 - max(min(tau.sq / tau.sq_full, 1.0), 0.0)) * 100.0
}
```

```
# Convenience settings for runjags in R:
#monitor# intercept, B_doorknock, B_SOGS, B_Years, B_Exp_EGM_pc_hdi, B_freq_thresh_m, B_freq_thresh_f, B_freq_thresh_w,
tau, tau.sq, tau_full, tau.sq_full, I.sq, H.sq, R.sq
#modules# glm on
#response# y
#residual# resid
#fitted# fitted_est


##########################################################################################################
##########################################################################################################
#### Initial values
##########################################################################################################
##########################################################################################################

inits{
"intercept" <- -1.5
"B_doorknock" <- -1.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- -0.5
}

inits{
"intercept" <- -0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- -1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
```

```
"B_Exp_EGM_pc_hdi" <- 1.5

}

inits{
"intercept" <- 0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- -1.5
}

inits{
"intercept" <- 1.5
"B_doorknock" <- -1.5
"B_SOGS"  <-  -0.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 1.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- 0.5
}
```

Listing G.3: Random effects meta-analysis of the prevalence of problem gambling, with moderators and fixed priors

```
model
      {
         for(i in 1:N) {
            y[i] ~ dbin(p[i], s_size[i])
            logit(p[i]) <- intercept +
                  B_doorknock * doorknock[i] +
                  B_SOGS * SOGS[i] +
                  B_freq_thresh_m * freq_thresh_m[i] +
                  B_freq_thresh_f * freq_thresh_f[i] +
                  B_freq_thresh_w * freq_thresh_w[i] +
                  B_Years * Years_before_2016[i] +
                  B_Exp_EGM_pc_hdi * Exp_TG_EGM_pc_hdi[i] +
                  theta[i]

            theta[i] ~ dnorm(0, prec)

            # Inverse variance weights for I2
            w[i] <- 1 / (1/y[i] + 1/(s_size[i] - y[i]))

            # Repeat with no covariates for R2 calc
            y_full[i] ~ dbin(p_full[i], s_size[i])
            logit(p_full[i]) <- intercept_full + theta_full[i]
            theta_full[i] ~ dnorm(0, prec_full)

            # calculate residuals and fitted values to bring back into R
            resid[i] <- y[i] - fitted_est[i]
            fitted_est[i] <- p[i] * s_size[i]
      }

      # Priors
      intercept ~ dnorm(0, 10^-6)
      B_doorknock ~ dnorm(0.0761, 10^24)
```

```
    B_SOGS ~ dnorm(0.4797, 10^24)
    B_freq_thresh_m ~ dnorm(-0.0332, 10^24)
    B_freq_thresh_f ~ dnorm( -0.0958, 10^24)
    B_freq_thresh_w ~ dnorm(-0.3965, 10^24)
    B_Years ~ dnorm(0, 10^-6)
    B_Exp_EGM_pc_hdi ~ dnorm(0, 10^-6)


    # Heterogeneity statistics
    N_params <- 8

    # Calculate tau and tau2
    prec <- 1/(tau.sq)
    tau.sq <- tau^2
    tau ~ dunif(0,10)  # Suggested by http://dx.doi.org/10.1002/sim.2112

    # Calculate I2
    # As defined on p. 1546, eq. 9 and 10 http://doi.org/10.1002/sim.1186
    H.sq <- (tau.sq + sigma.sq) / sigma.sq
    sigma.sq <- (sum(w) * (N - 1)) / (sum(w)^2 - sum(w^2))
    I.sq <- tau.sq / (tau.sq + sigma.sq)

    # R2 calc
    intercept_full ~ dnorm(0, 10^-6)
    tau_full ~ dunif(0,10)
    tau.sq_full <- tau_full^2
    prec_full <- 1/(tau.sq_full)
    R.sq <- (1 - max(min(tau.sq / tau.sq_full, 1.0), 0.0)) * 100.0
    }

# Convenience settings for runjags in R:
#monitor# intercept, B_doorknock, B_SOGS, B_Years, B_Exp_EGM_pc_hdi, B_freq_thresh_m, B_freq_thresh_f, B_freq_thresh_w,
tau, tau.sq, tau_full, tau.sq_full, I.sq, H.sq, R.sq
#modules# glm on
```

```
#response# y
#residual# resid
#fitted# fitted_est


######################################################################################################
######################################################################################################
#### Initial values
######################################################################################################
######################################################################################################

inits{
"intercept" <- -1.5
"B_doorknock" <- -1.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- -0.5
}

inits{
"intercept" <- -0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- -1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- 1.5

}

inits{
```

```
"intercept" <- 0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- -1.5
}

inits{
"intercept" <- 1.5
"B_doorknock" <- -1.5
"B_SOGS"  <-  -0.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 1.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- 0.5
}
```

Listing G.4: Random effects meta-analysis of the prevalence of moderate risk problem gambling, without moderators

```
model
      {
         for(i in 1:N) {
            y[i] ~ dbin(p[i], s_size[i])
            logit(p[i]) <- intercept + theta[i]

            # Calculate the weights to help with estimating Q
        w[i] <- 1 / (1/y[i] + 1/(s_size[i] - y[i]))

            theta[i] ~ dnorm(0, prec)
            resid[i] <- y[i] - fitted_est[i]
            fitted_est[i] <- p[i] * s_size[i]
      }

      # Calculate Q from Borenstein et al, p. 114, eq. 16.5
      N_params <- 1
      #Q <- (N - N_params) + C * tau.sq

      # Calculate Q from eq 16.1 on p. 109, Borenstein et al.
      Q <- sum(w * ((logit(p) - intercept)^2))

      # Calculate I2
      # As defined on p. 1546, eq. 9 and 10 http://doi.org/10.1002/sim.1186
      H.sq <- Q / (N - N_params)
      # I.sq.b <- min((H.sq - 1.0) / max(H.sq, 10^-12), 1.0) * 100

      # As defined on Borenstein et al., p. 117, eq. 16.9
      I.sq <- min((Q - (N - N_params)) / max(Q, 10^-12), 1.0) * 100.0

      intercept ~ dnorm(0, 10^-6)

      tau~  dunif(0,10)  # Suggested by http://dx.doi.org/10.1002/sim.2112
```

```
      tau.sq <- tau*tau
      prec <- 1/(tau.sq)
      }

# Convenience settings for runjags in R:
#monitor# intercept, tau, tau.sq, Q, I.sq, H.sq
#modules# glm on
#response# y
#residual# resid
#fitted# fitted_est

##########################################################################################################
##########################################################################################################
#### Initial values
##########################################################################################################
##########################################################################################################

inits{
"intercept" <- -1.5
}

inits{
"intercept" <- -0.5
}

inits{
"intercept" <- 0.5
}

inits{
"intercept" <- 1.5
}
```

Listing G.5: Random effects meta-analysis of the prevalence of moderate risk problem gambling, with moderators and informative priors

```
model
      {

         for(i in 1:N) {
            y[i] ~ dbin(p[i], s_size[i])
            logit(p[i]) <- intercept +
                  B_doorknock * doorknock[i] +
                  B_SOGS * SOGS[i] +
                  B_freq_thresh_m * freq_thresh_m[i] +
                  B_freq_thresh_f * freq_thresh_f[i] +
                  B_freq_thresh_w * freq_thresh_w[i] +
                  B_Years * Years_before_2016[i] +
                  B_Exp_EGM_pc_hdi * Exp_TG_EGM_pc_hdi[i] +
                  theta[i]

            theta[i] ~ dnorm(0, prec)

            # Inverse variance weights for I2
            w[i] <- 1 / (1/y[i] + 1/(s_size[i] - y[i]))

            # Repeat with no covariates for R2 calc
            y_full[i] ~ dbin(p_full[i], s_size[i])
            logit(p_full[i]) <- intercept_full + theta_full[i]
            theta_full[i] ~ dnorm(0, prec_full)

            # calculate residuals and fitted values to bring back into R
            resid[i] <- y[i] - fitted_est[i]
            fitted_est[i] <- p[i] * s_size[i]
      }

      # Priors
```

```
FLATTEN_FACTOR <- 4

intercept ~ dnorm(0, 10^-6)
B_doorknock ~ dnorm(0.7269,(1/(0.1780*FLATTEN_FACTOR))^2)
B_SOGS ~ dnorm(-0.5324, (1/(0.1395*FLATTEN_FACTOR))^2)
B_freq_thresh_m ~ dnorm(-0.1447, (1/(0.0579*FLATTEN_FACTOR))^2)
B_freq_thresh_f ~ dnorm( -0.3809, (1/(0.0559*FLATTEN_FACTOR))^2)
B_freq_thresh_w ~ dnorm(-0.5608, (1/(0.0610*FLATTEN_FACTOR))^2)
B_Years ~ dnorm(0, 10^-6)
B_Exp_EGM_pc_hdi ~ dnorm(0, 10^-6)


# Heterogeneity statistics
N_params <- 8

# Calculate tau and tau2
prec <- 1/(tau.sq)
tau.sq <- tau^2
tau ~ dunif(0,10)  # Suggested by http://dx.doi.org/10.1002/sim.2112

# Calculate I2
# As defined on p. 1546, eq. 9 and 10 http://doi.org/10.1002/sim.1186
H.sq <- (tau.sq + sigma.sq) / sigma.sq
sigma.sq <- (sum(w) * (N - 1)) / (sum(w)^2 - sum(w^2))
I.sq <- tau.sq / (tau.sq + sigma.sq)

# R2 calc
intercept_full ~ dnorm(0, 10^-6)
tau_full ~ dunif(0,10)
tau.sq_full <- tau_full^2
prec_full <- 1/(tau.sq_full)
R.sq <- (1 - max(min(tau.sq / tau.sq_full, 1.0), 0.0)) * 100.0
}
```

```
# Convenience settings for runjags in R:
#monitor# intercept, B_doorknock, B_SOGS, B_Years, B_Exp_EGM_pc_hdi, B_freq_thresh_m, B_freq_thresh_f, B_freq_thresh_w,
tau, tau.sq, tau_full, tau.sq_full, I.sq, H.sq, R.sq
#modules# glm on
#response# y
#residual# resid
#fitted# fitted_est


#######################################################################################################
#######################################################################################################
#### Initial values
#######################################################################################################
#######################################################################################################

inits{
"intercept" <- -1.5
"B_doorknock" <- -1.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- -0.5
}

inits{
"intercept" <- -0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- -1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- 1.5
```

```
}

inits{
"intercept" <- 0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- -1.5
}

inits{
"intercept" <- 1.5
"B_doorknock" <- -1.5
"B_SOGS"  <-  -0.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 1.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- 0.5
}
```

Listing G.6: Random effects meta-analysis of the prevalence of moderate risk problem gambling, with moderators and fixed priors

```
model
      {

         for(i in 1:N) {
            y[i] ~ dbin(p[i], s_size[i])
            logit(p[i]) <- intercept +
                  B_doorknock * doorknock[i] +
                  B_SOGS * SOGS[i] +
                  B_freq_thresh_m * freq_thresh_m[i] +
                  B_freq_thresh_f * freq_thresh_f[i] +
                  B_freq_thresh_w * freq_thresh_w[i] +
                  B_Years * Years_before_2016[i] +
                  B_Exp_EGM_pc_hdi * Exp_TG_EGM_pc_hdi[i] +
                  theta[i]

            theta[i] ~ dnorm(0, prec)

            # Inverse variance weights for I2
            w[i] <- 1 / (1/y[i] + 1/(s_size[i] - y[i]))

            # Repeat with no covariates for R2 calc
            y_full[i] ~ dbin(p_full[i], s_size[i])
            logit(p_full[i]) <- intercept_full + theta_full[i]
            theta_full[i] ~ dnorm(0, prec_full)

            # calculate residuals and fitted values to bring back into R
            resid[i] <- y[i] - fitted_est[i]
            fitted_est[i] <- p[i] * s_size[i]
      }

      # Priors
      intercept ~ dnorm(0, 10^-6)
```

```
    B_doorknock ~ dnorm(0.7269, 10^24)
    B_SOGS ~ dnorm(-0.5324, 10^24)
    B_freq_thresh_m ~ dnorm(-0.1447, 10^24)
    B_freq_thresh_f ~ dnorm(-0.3809, 10^24)
    B_freq_thresh_w ~ dnorm(-0.5608, 10^24)
    B_Years ~ dnorm(0, 10^-6)
    B_Exp_EGM_pc_hdi ~ dnorm(0, 10^-6)


    # Heterogeneity statistics
    N_params <- 8

    # Calculate tau and tau2
    prec <- 1/(tau.sq)
    tau.sq <- tau^2
    tau ~ dunif(0,10)  # Suggested by http://dx.doi.org/10.1002/sim.2112

    # Calculate I2
    # As defined on p. 1546, eq. 9 and 10 in http://doi.org/10.1002/sim.1186
    H.sq <- (tau.sq + sigma.sq) / sigma.sq
    sigma.sq <- (sum(w) * (N - 1)) / (sum(w)^2 - sum(w^2))
    I.sq <- tau.sq / (tau.sq + sigma.sq)

    # R2 calc
    intercept_full ~ dnorm(0, 10^-6)
    tau_full ~ dunif(0,10)
    tau.sq_full <- tau_full^2
    prec_full <- 1/(tau.sq_full)
    R.sq <- (1 - max(min(tau.sq / tau.sq_full, 1.0), 0.0)) * 100.0
    }

# Convenience settings for runjags in R:
#monitor# intercept, B_doorknock, B_SOGS, B_Years, B_Exp_EGM_pc_hdi, B_freq_thresh_m, B_freq_thresh_f, B_freq_thresh_w,
tau, tau.sq, tau_full, tau.sq_full, I.sq, H.sq, R.sq
```

```
#modules# glm on
#response# y
#residual# resid
#fitted# fitted_est


#######################################################################################################
#######################################################################################################
#### Initial values
#######################################################################################################
#######################################################################################################

inits{
"intercept" <- -1.5
"B_doorknock" <- -1.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- -0.5
}

inits{
"intercept" <- -0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- -1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- 1.5

}
```

```
inits{
"intercept" <- 0.5
"B_doorknock" <- -0.5
"B_SOGS"  <- 1.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 0.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- 1.5
"B_Exp_EGM_pc_hdi" <- -1.5
}

inits{
"intercept" <- 1.5
"B_doorknock" <- -1.5
"B_SOGS"  <-  -0.5
"B_freq_thresh_m"  <- 0
"B_freq_thresh_f"  <- 1.5
"B_freq_thresh_w"  <- 0
"B_Years"  <- -0.5
"B_Exp_EGM_pc_hdi" <- 0.5
}
```

Appendix H: Supplementary tables for Chapter 7

Table H.1: Pairwise Pearson's correlation coefficients for variables in Models 1 – 4. 2014 data only

	Incidents per 10,000	Assaults per 10,000	*ln* (Venues per 100,000 + 1)	*ln* (E.G.M.s per 10,000 + 1)	I.E.R.	I.E.R.^{2}	*ln* Fem. income share	*ln* % English only	Child-to-woman ratio	*ln* (% Indig. + 1)	Median age	Media age^2	*ln* (ARIA + 1)
Incidents.per.10000	1.00												
Assaults per 10,000	0.74	1.00											
ln (Venues per 100,000 + 1)	0.23	0.13	1.00										
ln (E.G.M.s per 10,000 + 1)	0.19	0.10	0.96	1.00									
I.E.R.	-0.41	-0.30	-0.02	0.01	1.00								
I.E.R.^{2}	0.04	-0.03	0.13	0.14	0.38	1.00							
ln Fem. income share	0.13	0.10	0.08	0.07	-0.12	0.12	1.00						
ln % English only	-0.08	-0.05	-0.33	-0.39	-0.12	-0.29	-0.04	1.00					
Child-to-woman ratio	0.06	0.05	-0.13	-0.15	-0.31	-0.37	-0.03	0.34	1.00				
ln (% Indigenous + 1)	0.29	0.20	0.01	-0.03	-0.44	-0.08	0.23	0.21	0.23	1.00			
Median age	-0.03	-0.02	-0.30	-0.34	-0.17	-0.23	0.21	0.48	0.23	0.11	1.00		
Median age^2	0.07	-0.02	-0.07	-0.06	0.07	0.18	0.06	-0.13	-0.12	-0.02	0.16	1.00	
ln (OARIA + 1)	-0.01	0.08	-0.34	-0.39	-0.32	-0.30	0.24	0.48	0.41	0.35	0.54	-0.01	1.00

Table H.2: Bivariate associations between domestic violence, EGM accessibility and socio-demographic characteristics from Bayesian spatio-temporal analysis

	Family incidents Unadjusted β coefficients		Domestic-violence assaults Unadjusted β coefficients	
	Est.	95% C.I.	Est.	95% C.I.
Intercept				
ln (Venues per 100,000 + 1) × 10^1	**1.4**	**[1.3, 1.6]**	**1.5**	**[1.3, 1.8]**
ln (E.G.M.s per 10,000 + 1) × 10^1	**1.0**	**[0.9, 1.1]**	**1.1**	**[0.9, 1.3]**
I.E.R. × 10^3	**-6.8**	**[-7.2, -6.3]**	**-7.3**	**[-7.9, -6.8]**
I.E.R.^{2} × 10^1	-1.6	[-5.3, 1.8]	3.1	[-2.0, 8.0]
ln Fem. income share	**1.7**	**[1.5, 2.0]**	**2.1**	**[1.7, 2.4]**
ln % English only × 10^1	**-5.9**	**[-7.6, -4.4]**	**-7.0**	**[-8.9, -4.7]**
Child-to-woman ratio × 10^1	3.4	[-0.1, 6.7]	5.2	[-0.1, 10.1]
ln (% Indigenous + 1) × 10^1	**5.5**	**[4.8, 6.3]**	**6.0**	**[5.1, 7.0]**
Median age × 10^2	**-1.2**	**[-1.9, -0.5]**	**-1.3**	**[-2.2, -0.3]**
Median age^2	4.6	[-2.8, 11.2]	5.2	[-2.8, 13.5]
ln (OARIA + 1) × 10^1	**-5.3**	**[-8.3, -2.2]**	**-3.9**	**[-8.3, 0.0]**

Notes: Est. = estimate; C.I. = credible interval; E.G.M. = electronic gaming machine; I.E.R. = Index of economic resources; Fem. = female; D.I.C. = Deviance information criterion. Bold type indicates coefficients for which the 95% C.I. does not contain zero. Unadjusted coefficients are estimates from a series of bivariate spatio-temporal models. Parameters were estimated simultaneously with their quadratic term in unadjusted models.