Impact of economic conditions and marriage market on inmates' recidivism: a longitudinal analysis

Mémoire

William Arbour

Maîtrise en économique - avec mémoire
Maître ès arts (M.A.)

Québec, Canada

Impact of Economic Conditions and Marriage Market on Inmates' Recidivism: a Longitudinal Analysis

Mémoire

William Arbour

Sous la direction de:

Guy Lacroix, directeur de recherche
Vincent Boucher, codirecteur de recherche

Résumé

Dans ce mémoire, nous étudions la population judiciarisée masculine québécoise en mettant l'accent sur la récidive. Plus précisément, nous cherchons à savoir si les conditions économiques et démographiques régionales dans lesquelles un individu est libéré affectent sa probabilité de récidive ainsi que le nombre de récidives commises au courant d'une année. À cet égard, nous étudions l'impact du taux de chômage, du salaire minimum et du revenu disponible moyen régional ainsi que le fait d'être marié ou d'avoir une famille. Comme le fait d'être marié et d'avoir une famille sont deux variables potentiellement endogènes avec la récidive, nous les instrumentons par la répartition régionale des sexes. À partir de données fournies par le ministère de la Sécurité publique, nous construisons un modèle de récidive en utilisant des régressions binomiales à variables instrumentales, que nous complétons avec un modèle de comptage de type Poisson, aussi estimé avec des variables instrumentales. Même si aucun effet n'est discernable quant au mariage, nous trouvons que d'avoir une famille diminue de 18% la probabilité de récidiver. Nous montrons que le mariage ou la famille n'ont pas d'effet sur le nombre de crimes commis par année, une fois que l'endogénéité est prise en compte. Finalement, nous mettons en lumière que le salaire minimum est négativement corrélé avec la récidive, alors que la richesse régionale moyenne semble être positivement corrélée avec la récidive.

Abstract

In this thesis, we study Quebec male prison population with a special focus on recidivism. We are especially interested in quantifying the impact of regional economic variables on the probability of recidivism and the number of reoffenses committed in a year. More precisely, we consider variations in regional unemployment rate, minimum wage and average disposable income. Moreover, we include in our analysis domestic predictors: being married or having a family. Since these two parameters are potentially endogenous with recidivism, we estimate bivariate probit regressions and Poisson regressions with regional sex ratios as instrumental variables. The data provided by the ministère de la Sécurité publique allows us to bring to light several results. We estimate that having a family decreases by 18% the likelihood of recidivism, while the effect marriage is not significant. No effect is found for the number of reoffenses committed, neither for being married or having a family. Finally, we find a significant positive correlation between regional wealth and recidivism, whereas we highlight that an increase in real minimum wage is associated with a decrease in recidivism.

Contents

List of Tables

List of Figures

À ma cousine, Nathalie.

The nuns taught us there are two ways through life: the way of Nature and the way of Grace. You have to choose which one you'll follow. Grace doesn't try to please itself... accepts being slighted, forgotten, disliked... accepts insults and injuries. Nature only wants to please itself... gets others to please it too... likes to lord it over them... to have its own way. It finds reasons to be unhappy... when all the world is shining around it... when love is smiling through all things.

— Terrence Malick, *The Tree of Life*

Remerciements

À celles et ceux qui seront mentionnés ici, sachez que je pèse la justesse de ces mots depuis un bon moment déjà. Ainsi, voyez-les comme l'expression la plus authentique de mes sentiments.

Au Professeur Guy Lacroix, qui m'a pris aveuglément sous son aile et qui m'a proposé un projet qui a su me fasciner pendant près d'une année. Monsieur Lacroix, votre abnégation dans toutes les sphères de votre travail m'a permis de devenir un meilleur chercheur, mais m'a aussi montré l'étendue des efforts à déployer pour devenir un scientifique doté de votre rigueur. À la lecture de ce mémoire, j'espère que vous ressentirez qu'une infime partie de la fierté que j'ai eue à l'écrire sous votre direction. Merci.

Au Professeur Vincent Boucher, dont la passion contagieuse pour la recherche m'a poussé à élever le niveau de ce mémoire plus haut que je ne l'aurais d'abord espéré. Vincent, ce serait un honneur de retravailler avec toi dans le futur. Merci.

Je ne saurais passer sous silence l'apport considérable qu'a eu Nicolas Corneau-Tremblay durant mes études de maîtrise. Nicolas, tu es des gens dont la qualité exceptionnelle n'a d'égale que l'intelligence, et ce n'est pas peu dire dans ton cas. Si, un jour, je ne répète qu'une fraction de ce que tu as fait pour moi, je saurai que j'aurai réussi. D'ici là, bon succès dans les nouvelles aventures qui t'attendent. Merci.

À Steeve Marchand, dont les conseils éclairants se sont révélés essentiels à plusieurs reprises et dont la générosité est sans bornes. Merci. J'en profite pour saluer Laure Sébrier, qui a su m'écouter d'une oreille toujours attentive - parfois, l'écoute vaut plus que n'importe quelle parole. Merci. À vous deux, que l'expérience en Californie ne puisse que multiplier la passion qui vous anime déjà.

Au Professeur Markus Herrmann, pour son intelligence vive, son écoute, son écoute et son écoute. Quelquefois, c'est tout ce qu'il nous faut. Merci. Au Professeur Bernard Fortin, qui m'a transmis un peu de sa vaste culture (de Bergman à Truffaut, en passant par Fellini) et qui a su m'aiguiller sur plusieurs plans. Merci.

À mes *amis du bac*, qui ont été finalement bien plus que cela, Jean-Charles, Alexis et Étienne, qui sont responsables en partie de la poursuite de mes études en économie. Merci. À *mes* *amis tout court*, qui sont ma source de motivation et, de façon beaucoup plus importante, ma source première de fierté. À Camille, Andréanne, Maxence, Lorrie, David, Anne-Sophie, Laurence, William et Olivier. Vous contribuez, chacun à votre façon, à faire de moi une meilleure personne. Merci.

Je souhaite remercier Monsieur Bernard Chéné, du ministère de la Sécurité publique, dont les réponses à mes questions ont été indispensables. Merci.

À mes parents, Michèle et François. Il serait injuste de vous écrire un simple *merci* comme je l'ai fait plus haut. J'espère qu'en disant que je ne pourrais pas me sentir plus aimé, vous comprendrez l'ampleur que cela représente.

Chapter 1

Introduction

In the province of Quebec, the ministère de la Sécurité publique (MSP) is in charge of the application of the Act Respecting the Quebec Correctional System, which purpose is to encourage reintegration of offenders. Although Quebec detention centers record more than 10,000 passages each year, the Report of the Auditor General of Quebec to the National Assembly for 2016-2017 highlighted a few issues in the correctional services. Among other conclusions, we can read that MSP "does not know the result of its reintegration efforts". In fact, no academic studies have examined the male prison population of Quebec with a focus on recidivism. To overcome this vagueness in the literature and in order to gain a better understanding of recidivism among Quebec offenders, MSP provided us with an exclusive dataset.

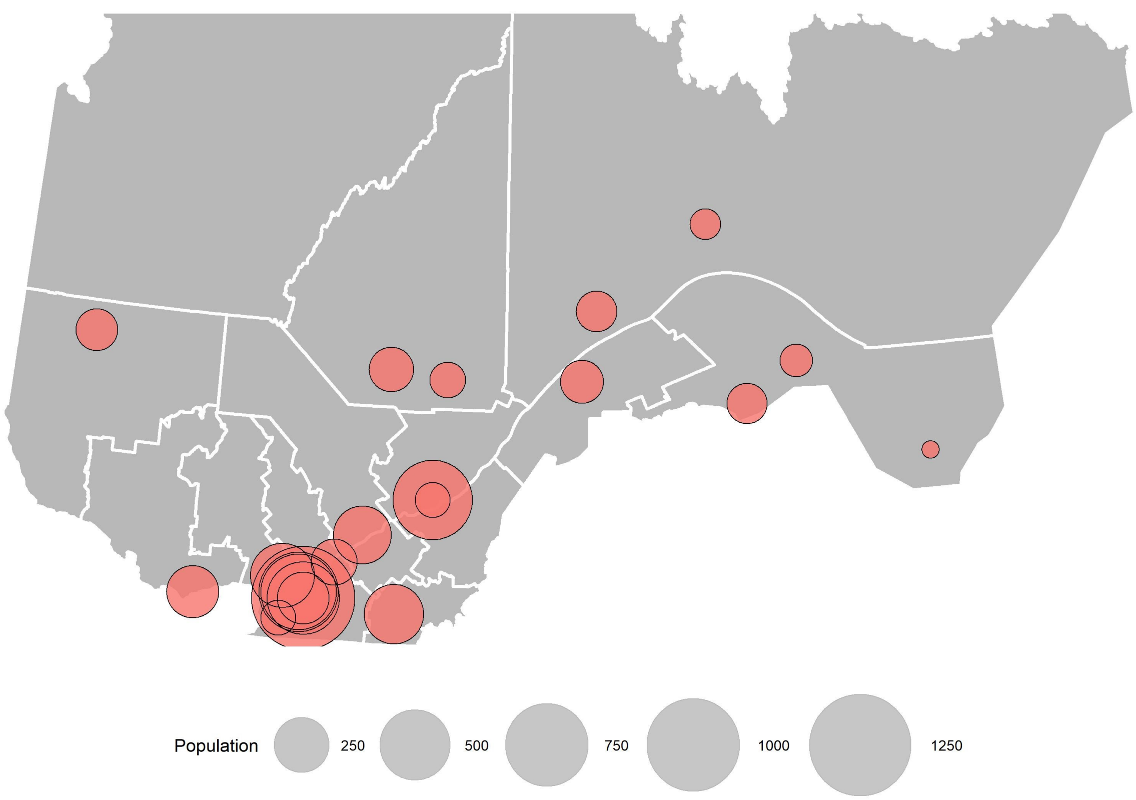
Before moving forward, it seems essential to put our issue against the peculiar context of the Quebec Judicial System. In 1867, The Constitution Act distributed the legislative powers: Article 91(28) states that the Government of Canada is responsible for the management of penitentiaries, whereas Article 92(6) specifies that public prisons are under the provincial legislature. Since then, sentences of less than two years are served under each province responsibility. Therefore, in Quebec, MSP's legislature concentrates primarily on the sentences associated with possession of substance, break-ins, theft and failure to comply with condition of undertaking or recognizance. We present, in Table 1.1, the five most common offenses committed between 2007 and 2017, where the listed codes refer to either Criminal Code (R.S.C., 1985, c. C-46) or Controlled Drugs and Substances Act (S.C. 1996, c. 19).

Table 1.1 – Most Common Offenses Among Men and Women Over 2007-2017

	Male			Female		
	Code	Nb. of Admissions	Offense	Code	Nb. of Admissions	Offense
1	CC0145	94436	Failure to attend court Failure to comply with condition of undertaking or recognizance	CC0145	10146	Failure to attend court Failure to comply with condition of undertaking or recognizance
2	CC0733	70960	Failure to comply with probation order	CC00733	8908	Failure to comply with probation order
3	CC0334	36945	Theft Forgery	CC0334	5011	Theft Forgery
4	CC0264	31677	Criminal harassment Uttering threats	ST0004	2440	Possession of substance
5	ST0004	29401	Possession of substance	CC0264	2211	Criminal harassment Uttering threats

When an individual is found guilty of an offense (or several offenses) and if the sentence to be served is less than two years, the individual is sent to one of the 20 detention centers, as displayed on Figure 1.1, in which the size of each circle is proportional to the population the center can accommodate. In 2013-2014, the real capacity was 4,900 inmates, while recent developments made it possible to take in around 1,000 more inmates from 2016.

Figure 1.1 – Geographic Location and Capacity of Quebec Detention Centers



Among these centers, Établissement de détention de Percé is specialized in sexual delinquency and Établissement de détention d'Amos has reserved areas for Inuits and First Nations. Specific programs are offered in both of these centers. All of the women inmates are sent to Établissement de détention Leclerc[1] or in Établissement de détention de Québec, which totals approximately 300 places for women. Even though Maison Tanguay and Établissement de détention de Chicoutimi both closed in 2015, we still considered these centers in our analysis as our data documents the passages in these centers as well.

In this thesis, we are interested in the impact of regional economic conditions, such as the regional unemployment rate and the average disposable income, and the competition on the marriage market on recidivism. Mainly, our research question is twofold:

- Given personal characteristics, how do regional economic conditions and marriage market impact the **probability** of recidivism?
- Given personal characteristics, how do regional economic conditions and marriage market impact the **frequency** of recidivism on an annual basis?

The pertinence of our contribution lies in the fact that the Quebec male prison population has never been the subject of econometric studies before. Furthermore, we take into account the endogeneity of matrimonial status and the fact of having a family, and we take advantage of regional sex ratios to isolate a causal effect of these parameters on recidivism. Our goal is to answer our research questions using rigorous econometric methods such as probit modeling, instrumental variables and regression analysis of count data but also to provide a first exploratory analysis that will pave the way to a series of original research questions.

A difficulty one encounters while studying recidivism is the lack of unanimity surrounding the very definition of recidivism. In this thesis, we consider recidivism as the simple fact of returning to jail without regards to the type of offenses committed nor the time elapsed between crimes.

The paper is organized as follows. Chapter 2 starts by presenting an overview of the existing literature about the predictors of recidivism and provides a special focus on economic conditions. Chapter 3 explains the methodology used and the estimation methods. Chapter 4 documents the data sources, defines the hypotheses being made and presents descriptive statistics of the Quebec prison population. Chapter 5 exhibits the results to our research questions. We finally briefly conclude by identifying the limitations of our study and ideas for future research.

[1]Since Maison Tanguay closed in 2015. Thus, Établissement de détention Leclerc became a mixed prison.

Chapter 2

Literature Review

On n'empêche pas plus la pensée de revenir à une idée que la mer de revenir à un rivage. Pour le matelot, cela s'appelle la marée; pour le coupable, cela s'appelle le remords.

— Victor Hugo, *Les Misérables*

In a leading article, Becker (1968) argues that criminals evaluate the costs and the benefits of a crime before committing it, implying that criminals act rationally. Becker suggests that an individual chooses to commit a crime if his expected utility is greater than his expected utility of not committing a crime. That is, as Chalfin and McCrary (2017) put it, an individual will take action if

$$\mathbb{E}(U) = pU(Y - f) + (1 - p)U(Y) > U_{nc}, \tag{2.1}$$

where p is the probability of getting caught, Y stands for the benefits of the crime and f is the severity of the sanction. Finally, U_{nc} represents the utility of abstaining from crime.

Since then, sociologists, criminologists and economists have tried to model and predict crime using different variables that will affect either p, Y or f; a large part of the research being to determine what deters a potential criminal to take action and which risk factors will have an impact on recidivism.

Over the years, research has highlighted several predictors of recidivism. For Andrews and Bonta (2010), the predictors of criminal behaviour can be categorized as static or dynamic. Static risk factors do not vary over time and include variables such as sex, being young, the presence of a criminal record and having been raised in a family getting social assistance. Dynamic predictors can change over time and may be geographic position, employment, marriage and social network, to name a few. According to Gendreau et al. (1996), a scientific consensus exists surrounding the importance of various static predictors, but some divergence persists on dynamic factors because of ideological concerns in the scientific community and methodological difficulties. Note that age is sometimes considered as static by certain researchers and dynamic by others. In this literature review, we will explore recent work paying attention to both static and dynamic predictors as well as to papers dealing with economic conditions and marriage market.

2.1 Static Predictors

In a meta-analysis on 131 studies on Canadian and American data published between 1970 and 1994, Gendreau et al. (1996) show that the most significant correlated covariates among the studies are age, gender, criminal history, family criminality, socioeconomic status, family structure, race and family rearing practices.

Literature comparing men's predictors to women's is not extensive. Nevertheless, Benda (2005) proves, using a Cox proportional hazard model on 300 women and 300 men, that predictors such as childhood, recent abuse, the presence of a criminal partner, depression, selling drugs and suicidal thoughts are stronger for women than for men. The same discussion is made by Bonta et al. (1995), who insists on variables that usually do not predict recidivism among men but that have a positive correlation with women's: history of juvenile delinquency, crime committed with an associate, alcohol and drug abuse. Another interesting difference between men and women is brought up by Mitchell et al. (2017). With a regression discontinuity design, they show that prison had a criminogenic effect on men, but not on women. They also argue that prison, in comparison to non-incarcerative sanctions, had no significant effect on recidivism. A different conclusion is reached by Drago et al. (2009). Indeed, they support that spending an additional month in prison reduces the probability of recidivism by 1.24%, depending on the time first served incarcerated.

About race, Jung et al. (2010) use survival analysis on 12,545 inmates from Pennsylvania to show that Afro-American men recidivate at a significantly higher rate than white men, with a shorter survival time, meaning that Afro-American men recidivate faster. The same conclusions were published by McGovern et al. (2009), who find, with a weighted logistic regression on 142,095 American individuals, that white men have a lower recidivism rate than Hispanic men, who in turn have a lower rate than black men.

On age, Olver and Wong (2015) and Piquero et al. (2015) conducted studies on Canadian and American data and both studies demonstrated that age was negatively correlated with recidivism, suggesting that young people tend to repeat an offense more than older people. Note that age is much more often studied with sexual offenders. As regards to this, see Hanson (2002), Thornton (2006) and Stephens et al. (2016).

Finally, it is well documented by Kulychek (in studies from 2006 and 2007) that offenders with a criminal record have a higher probability of committing a crime than people without records in their first adult years. At around 23 years of age, hazard rates for both groups converge and show little difference afterwards.

2.2 Dynamic Predictors

In their meta-analysis, Gendreau et al. (1996) also bring forth several significant dynamic predictors to light: antisocial personality, companions, criminogenic needs, interpersonal conflict, personal distress, social achievement and substance abuse. Recent research confirms these results but also suggests new significant explanatory variables.

Drago et al. (2011), and Chen and Shapiro (2007) evaluate the impact of conditions in prison on recidivism. Respectively on Italian and American data, they demonstrate that harshness in prison does not lead to the reduction of the propensity to commit another crime. In fact, harshness is just positively correlated with a return to prison.

On a social level, Bales and Mears (2008) use transition data analysis to study the impact of visitation during the incarceration on recidivism. The results indicate that visitation of family members is significant, with emphasis on spousal visitation, which has slightly more impact on reducing the likelihood of recidivism. Marriage is also an important factor at the time of release, as stated by Andersen et al. (2015). With propensity score matching on 102,839 inmates, their results prove that marriage to an unconvicted spouse reduces the probability of recidivism in comparison to a marriage to a spouse with a criminal record.

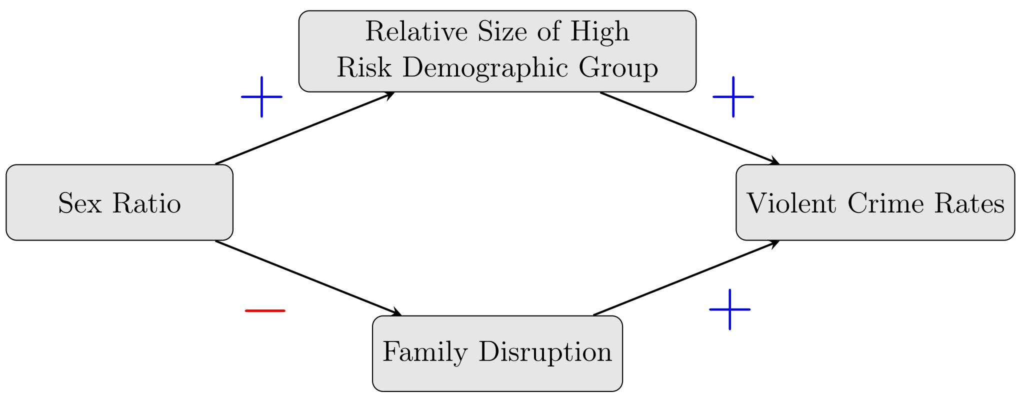
An important predictor appears to be employment after release. Denver et al. (2017) calculate that being employed reduces by 2.2% the probability of recidivism during the first year and by 4.2% for the next two years. Similar conclusions are found in research by Tripodi et al. (2010), by Visher et al. (2005) and more recently by Siwach (2018). In the latter, the author finds, using 2SLS estimator, that an additional quarter of unemployment increases by 6% the likelihood of being arrested again.

2.3 Economic Conditions

Although the literature studying individual characteristics of repeat offenders using micro data is broad, the literature about the impact of regional economic conditions on recidivism, on a macro level, is less extensive. Among the early few to study this issue is Freeman (1983), who analyzes how participation in crime may be induced by economic life. According to Freeman (1999), this literature lacks robustness: in spite of the fact that the elasticity of the supply of crime is expected to be high, meaning that potential offenders might trade legal labour for crime, they might not be that responsive to aggregate data of economic conditions. However, recent investigations turned out to be more conclusive. Following this example, Corman and Mocan (2005) use the unemployment rate and the minimum wage to measure economic conditions: a significant impact is found for all types of felonies, except for assault and rape. In regard to unemployment, Lin (2008) finds an elasticity of 4% on crime rates, while Fernandez et al. (2014) conclude, using various regression models on American data, that minimum wage is slightly correlated with reduction of property crimes[1].

Few articles discuss the effect of the regional sex ratio on crime and, even fewer so, on recidivism. However, according to Messner and Sampson (1991), sex ratio[2] could be an important factor for it creates competition on the marriage market, a hypothesis recently supported by Rocque et al. (2015). Furthermore, the effect could be spread by two main channels, as displayed on Figure 2.1.

Figure 2.1 – Transmission Effect of Sex Ratio on Violent Crime Rates



The intuition behind Figure 2.1 goes as follows: a higher sex ratio (more men relatively to women) implies a larger number of people at risk since men tend to participate in crime activities more than women do. Also, Messner and Sampson (1991) suggests that a higher sex ratio leads to fewer family disruptions, a factor that would generally increase crime rates. Since the effects cancel each other out, we expect the overall effect to be rather small. Nevertheless, Edlund et al. (2013) take advantage of China's one child policy and traditional son preference to discuss the implications of an increasing sex ratio. They estimate an ordinary least squares regression on a 1988-2004 panel data and find that an increase of 1% in the sex ratio rises the likelihood of violent and property crime by around 3.5%.

2.4 Contribution

Our main contribution lies in the fact that Quebec's prison population has never been studied with so many predictors in an econometric research. Moreover, although a few studies attempt to explain how economic conditions affect crime rates, it is yet to be determined whether the unemployment rate, the minimum wage, the regional average disposable income and sex ratios

[1]No effect is found for non-property crimes (assault, rape, etc.).
[2]Sex ratio is calculated as the number of men divided by the number of women.

can have a recurrent effect on an individual, that is whether these variables are involved with recidivism.

Chapter 3

Methodology

Parents know when their children are pretending to be asleep.

— Paolo Sorrentino, *Youth*

3.1 Choice of Instrument

In the linear case, for instance $y = \mathbf{x}'\beta + \mu$, it can be easily shown that endogenous regressors ($\mathbb{E}(\mu|\mathbf{x}) \neq 0$) lead to inconsistent estimators. A method one can use to overcome this hurdle is the instrumental variables estimation. If we have $\mathbf{z}$ such that $\mathbb{E}(\mu|\mathbf{z}) = 0$ and if $\mathbf{x}$ and $\mathbf{z}$ are correlated, we call $\mathbf{z}$ the instrumental variables for endogenous variables $\mathbf{x}$.

Our analysis will be divided into two parts. In addition to taking into account classic economic conditions such as the minimum wage, the unemployment rate and the disposable income, we will test the impact of being married and the impact of having a family on recidivism. These two variables are defined below and will be extensively referred to afterwards.

$$marriage = \begin{cases} 1 \text{ if the individual is married;} \\ 0 \text{ otherwise.} \end{cases} \qquad family = \begin{cases} 1 \text{ if the number of dependants } > 0; \\ 0 \text{ otherwise.} \end{cases}$$

We will also define the variable $recidivism_{ik}$ such as

$$recidivism_{ik} = \begin{cases} 1 \text{ if individual } i \text{ reoffends in year } k; \\ 0 \text{ otherwise.} \end{cases}$$

We have reasons to believe that $marriage$ and $family$ could be endogenous with $recidivism$. First, endogeneity could result from omitted variables that would be correlated to both $marriage$ (or $family$) and $recidivism$, such as personal characteristics bound up with the individual's personality, such as religious beliefs. Second, we could imagine a structural model in which both variables affect each other (taking *marriage* as an example):

$$marriage = f(recidivism, \dots); \tag{3.1}$$

$$recidivism = f(marriage, \dots). \tag{3.2}$$

Although we wish to estimate Equation 3.2, each term appears in Equation 3.1, which results in a double causality. One could challenge the existence of Equation 3.1, stating that our population of interest is already criminal and that therefore recidivism does not influence the probability of getting married. We have to recall the scale of the crimes committed by our population: the majority of the sentences are short and are the consequences of minor crimes, and thus one offense probably does not deter potential partners. On the other hand, a reoffense could be the breaking point in the probability for an ex-offender to find someone to share his life with. Also, the time served in jail necessarily results in less time spent to find a partner. The same discussion applies to $family$. The appropriate tests will be presented in Chapter 5.

The chosen instrument is regional sex ratios, for which the calculation method is detailed at Section 4.2.3. We believe that sex ratios are a good instrument for two reasons:

1. Sex ratios are highly correlated with the probability of getting married and the probability of founding a family, as evidenced by, *inter alia*, Angrist (2002), Chiappori et al. (2002) and South and Lloyd (1992);
2. The only way sex ratios could have an impact on recidivism is by the transmission of marriage and family, as previously displayed in Figure 2.1.

Notice that our analysis will use either *marriage* or *family*, never both variables, since we only have one instrument. Hence, our models will be just-identified. The next parts will go further into the estimation methods using instrumental variables.

3.2 Binary Regressions

3.2.1 Introduction

Binary regressions will be useful to predict recidivism since we wish to estimate

$$\mathbb{P}(recidivism \mid X),$$

where the variable *recidivism* could take only the value of 0 or 1, as defined above, and where X contains the explanatory covariates. These regressions will allow us to answer our first question: *given personal characteristics, how do regional economic conditions and marriage market impact the probability of recidivism?*

To gain a better understanding of binary regressions, let us define the following probability:

$$\mathbb{P}(Y = 1 \mid X) = G(X'\beta).$$

As detailed by Kennedy (2003), G can take several forms: for a probit regression, $G = \Phi$, the cumulative Gaussian distribution and, for a logit regression, $G(X'\beta) = \frac{e^{X'\beta}}{1+e^{X'\beta}}$. Both models can be estimated by maximum likelihood:

$$\mathcal{L} = \prod_i \mathbb{P}(Y_i = 1 \mid x_i)^{Y_i}(1 - \mathbb{P}(Y_i = 1 \mid x_i))^{1-Y_i}. \tag{3.3}$$

For example, for the probit model, Equation 3.3 can be expressed as

$$\ln \mathcal{L} = \sum_i (y_i \ln \Phi(x_i'\beta) + (1 - y_i)\ln(1 - \Phi(x_i'\beta))).$$

Then our estimator, $\hat{\beta}$, will be

$$\hat{\beta} = \arg\max_{\beta} \ln \mathcal{L}.$$

3.2.2 Bivariate Probit Regression

Remember that we suspect our binary regressors, previously defined, to be endogenous. In this case, Baum et al. (2012) suggest estimating a bivariate probit model. In this section, we will describe this model and its estimation by maximum likelihood. Our presentation follows Greene (2003), and Cameron and Trivedi (2005).

Consider the following model, with unobserved latent variables y_1* and y_2*:

$$y_1* = \mathbf{x}_1'\beta_1 + \epsilon_1$$
$$y_2* = \mathbf{x}_2'\beta_2 + \epsilon_2,$$

in which

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).$$

We observe outcomes y_1 and y_2:

$$y_1 = \begin{cases} 1 \text{ if } y_1* > 0; \\ 0 \text{ otherwise.} \end{cases}$$

$$y_2 = \begin{cases} 1 \text{ if } y_2* > 0; \\ 0 \text{ otherwise.} \end{cases}$$

In our case, we have $y_1 = recidivism$, $y_2 = marriage$ (or $family$) and $\mathbf{x}_2'$ will be constructed as the augmented matrix $\mathbf{x}_2' = [\mathbf{x}_1' \ \textit{sex ratio}]$. In order to estimate β_1, β_2 and ρ, one must maximize the likelihood of the joint bivariate model. We must consider each of the following probability:

$$\begin{aligned} Q_{11} &= \mathbb{P}(y_1 = 1, y_2 = 1 | \beta_1, \beta_2)^{y_1 y_2}; \\ Q_{01} &= \mathbb{P}(y_1 = 0, y_2 = 1 | \beta_1, \beta_2)^{(1-y_1)y_2}; \\ Q_{10} &= \mathbb{P}(y_1 = 1, y_2 = 0 | \beta_1, \beta_2)^{y_1(1-y_2)}; \\ Q_{00} &= \mathbb{P}(y_1 = 0, y_2 = 0 | \beta_1, \beta_2)^{(1-y_1)(1-y_2)}. \end{aligned}$$

Substituting for the latent variables and taking the log, these quantities can be rewritten as

$$\begin{aligned} Q_{11} &= y_1 y_2 \log \mathbb{P}(\epsilon_1 > -\mathbf{x}_1'\beta, \epsilon_2 > -\mathbf{x}_2'\beta_2) = y_1 y_2 \log \Phi(\mathbf{x}_1'\beta_1, \mathbf{x}_2'\beta_2, \rho); \\ Q_{01} &= (1-y_1)y_2 \log \mathbb{P}(\epsilon_1 < -\mathbf{x}_1'\beta, \epsilon_2 > -\mathbf{x}_2'\beta_2) = (1-y_1)y_2 \log \Phi(-\mathbf{x}_1'\beta_1, \mathbf{x}_2'\beta_2, -\rho); \\ Q_{10} &= y_1(1-y_2) \log \mathbb{P}(\epsilon_1 > -\mathbf{x}_1'\beta, \epsilon_2 < -\mathbf{x}_2'\beta_2) = y_1(1-y_2) \log \Phi(\mathbf{x}_1'\beta_1, -\mathbf{x}_2'\beta_2, -\rho); \\ Q_{00} &= (1-y_1)(1-y_2) \log \mathbb{P}(\epsilon_1 < -\mathbf{x}_1'\beta, \epsilon_2 < -\mathbf{x}_2'\beta_2) = (1-y_1)(1-y_2) \log \Phi(-\mathbf{x}_1'\beta_1, -\mathbf{x}_2'\beta_2, \rho), \end{aligned}$$

where Φ is the cumulative function of a bivariate Gaussian distribution. The log-likelihood is therefore given by

$$\mathcal{L} = Q_{11} + Q_{01} + Q_{10} + Q_{00}$$

and the estimators will maximize this function as

$$\hat{\hat{\beta}} = [\hat{\beta}, \hat{\rho}] = \max_{\beta,\rho} \mathcal{L}.$$

As for any binary regression, $\hat{\beta}$ is not directly interpretable: marginal effects have to be computed afterwards. Later on in the analysis, a Wald test based on Wooldridge (2010) will be calculated to test for endogeneity, that is, if $\rho = 0$.

3.2.3 Probit Model with Continuous Endogenous Regressors

First, let's define the following model:

$$y_{1i}* = y_{2i}\beta + X_{1i}\gamma + \mu_i \tag{3.4}$$

$$y_{2i} = X_{1i}\Pi_i + x_{2i}\pi_2 + v_i, \tag{3.5}$$

where y_{2i} is the endogenous regressor (in our case, *marriage* or *family*), X_{1i} contains other explanatory covariates and x_{2i} is the instrument for y_{2i} (in our case, *sex ratio*). As for a regular probit model, we only observe

$$y_{1i} = \begin{cases} 0 \text{ if } y_{1i}* < 0; \\ 1 \text{ if } y_{1i}* \geq 0. \end{cases}$$

For us, $y_{1i} = 1$ if the i-th individual reoffended at least once during a precise year. Although μ_i and v_i are not independent, they are jointly normal:

$$\begin{pmatrix} \mu_i \\ v_i \end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & \rho\sigma_v \\ \rho\sigma_v & \sigma_v^2 \end{bmatrix}\right),$$

since we imposed $\mathbb{V}(\mu_i) = \sigma_\mu = 1$ as in any probit model. We then can express the conditional

$$\mu_i | v_i \sim \mathcal{N}(\rho\sigma_v v_i, \sigma_v^2(1-\rho^2)).$$

Therefore, $\mu_i = \rho\sigma_v v_i + \eta$, where $\eta_i \sim \mathcal{N}(0, \sigma_v^2(1-\rho^2))$. Equation 3.4 can be rewritten:

$$\begin{aligned} y_{1i}* &= y_{2i}\beta + X_{1i}\gamma + \mu_i \\ &= y_{2i}\beta + X_{1i}\gamma + \rho\sigma_v v_i + \eta_i \\ &= y_{2i}\beta + X_{1i}\gamma + \rho\sigma_v(y_{2i} - X_{1i}\Pi_1 - x_{2i}\pi_2) + \eta. \end{aligned} \tag{3.6}$$

The joint density for observation i is given by $f(y_{1i}, y_{2i}|x_i)$, where $x_i = [X_{1i}, x_{2i}]$. We can decompose this density with the conditional density rule:

$$f(y_{1i}, y_{2i}|x) = f(y_{1i}|y_{2i}, x_i)f(y_{2i}|x_i).$$

We will derive each density separately. For notation matters, let

$$\kappa_i = y_{2i}\beta + X_{1i}\gamma + \rho\sigma_v(y_{2i} - X_{1i}\Pi_1 - x_{2i}\pi_2),$$

the right-hand side of Equation 3.6 without the error term. We have

$$\begin{aligned} f(y_{1i}|y_{2i}, x_i) &= \mathbb{P}(y_{1i} = 1)^{y_{1i}}\mathbb{P}(y_{1i} = 0)^{1-y_{1i}} \\ &= \mathbb{P}(y_{1i}* \geq 0)^{y_{1i}}\mathbb{P}(y_{1i}* < 0)^{1-y_{1i}} \\ &= \mathbb{P}(\kappa_i + \eta_i \geq 0)^{y_{1i}}\mathbb{P}(\kappa_i + \eta_i < 0)^{1-y_{1i}} \\ &= \mathbb{P}(\eta_i \geq -\kappa_i)^{y_{1i}}\mathbb{P}(\eta_i < -\kappa_i)^{1-y_{1i}}. \end{aligned}$$

Recall that $\eta_i \sim \mathcal{N}(0, \sigma_v^2(1-\rho^2))$. We can now standardize η_i.

$$f(y_{1i}|y_{2i}, x_i) = \mathbb{P}\left(\frac{\eta_i}{\sqrt{\sigma_v^2(1-\rho^2)}} \geq -\frac{\kappa_i}{\sqrt{\sigma_v^2(1-\rho^2)}}\right)^{y_{1i}} \mathbb{P}\left(\frac{\eta_i}{\sqrt{\sigma_v^2(1-\rho^2)}} < -\frac{\kappa_i}{\sqrt{\sigma_v^2(1-\rho^2)}}\right)^{1-y_{1i}}.$$

Let $\tilde{\kappa}_i = \frac{\kappa_i}{\sqrt{\sigma_v^2(1-\rho^2)}}$. Taking advantage of the symmetry of the Gaussian distribution, the last equation can be expressed as

$$\begin{aligned} f(y_{1i}|y_{2i}, x_i) &= \left(1 - \mathbb{P}\left(\frac{\eta_i}{\sqrt{\sigma_v^2(1-\rho^2)}} \leq -\tilde{\kappa}_i\right)\right)^{y_{1i}} \mathbb{P}\left(\frac{\eta_i}{\sqrt{\sigma_v^2(1-\rho^2)}} < -\tilde{\kappa}_i\right)^{1-y_{1i}} \\ &= \Phi(\tilde{\kappa}_i)^{y_{1i}}(1 - \Phi(\tilde{\kappa}_i))^{1-y_{1i}}. \end{aligned} \tag{3.7}$$

This completes the derivation for $f(y_{1i}|y_{2i}, x_i)$. The derivation for $f(y_{2i}|x_i)$ is straightforward for it is a linear combination of a Gaussian distribution. Therefore,

$$f(y_{2i}|x_i) = \phi\left(\frac{y_{2i} - x_i}{\sigma_v}\right). \tag{3.8}$$

Equations 3.7 and 3.8 yield to the following likelihood:

$$\mathcal{L}_i = \prod_{i=1}^{N} \Phi(\tilde{\kappa}_i)^{y_{1i}}(1-\Phi(\tilde{\kappa}_i))^{1-y_{1i}}\phi\left(\frac{y_{2i}-x_i}{\sigma_v}\right).$$

Similarly as before, the estimated coefficients will not be directly interpretable: in our analysis, we will compute the marginal effects. As for the bivariate probit regression, a Wald test for endogeneity will be performed on $\hat{\rho}$, the estimated correlation parameter between μ_i and v_i, following Wooldridge (2010).

3.3 Count Models

3.3.1 Introduction

We now define the following variable:

$$\mathit{freq}_{ik} = \text{number of offenses committed by individual } i \text{ in year } k.$$

Since $\mathit{freq} \in \mathbb{N}$, we will use count models to predict $\mathbb{E}(\mathit{freq}|\mathbf{x})$, where $\mathbf{x}$ contains the explanatory variables. In other words, we will answer our second question: *given personal characteristics, how do regional economic conditions and marriage market impact the frequency of recidivism on an annual basis?*

Among other models, we will focus on the Poisson regression. The mass function of the Poisson distribution is given by

$$f(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}, \tag{3.9}$$

with $\mathbb{E}(y) = \mathbb{V}(y) = \lambda$. A Poisson regression occurs if we specify $\lambda = \exp(\mathbf{x}'\beta)$. Substituting into 3.9, the density for observation i is

$$f(y|\mathbf{x},\beta) = \frac{e^{-\exp(\mathbf{x}'\beta)}\exp(\mathbf{x}'\beta)^y}{y!},$$

where $y = \mathit{freq}$ in our case. Taking the product over all the sample $(i = 1,\dots,n)$ and the log lead to the following likelihood:

$$\mathcal{L} = \frac{1}{N}\sum_{i}^{N} -\exp(\mathbf{x}_i'\beta) + y_i\mathbf{x}_i'\beta - \log y_i!, \tag{3.10}$$

and $\hat{\beta}$ is such that

$$\hat{\beta} = \max_{\beta} \mathcal{L}.$$

Deriving Equation 3.10 with respect to β leads to the following nonlinear equations to solve:

$$\sum_{i}^{N}(y_i - \exp(\mathbf{x}_i'\beta))\mathbf{x}_i = \mathbf{0}. \tag{3.11}$$

The problem when one variable is endogenous is that $\mathbb{E}(y_i - \exp(x_i'\beta)|\mathbf{x}_i) \neq 0$ in a way that resolving the first-order conditions 3.11 yields to biased estimates.

3.3.2 Poisson Regression with Instrumental Variables

Since we suspect *marriage* and *family* to be endogenous, we will cover the estimation of a Poisson regression with instrumental variables. Let's suppose a multiplicative model in which y_{2i} is endogenous:

$$y_i = \exp\left(\mathbf{x}_i'\beta_1 + y_{2i}'\beta_2\right)\epsilon_i, \tag{3.12}$$

where ϵ_i is an error term. Cameron and Trivedi (2013) suggest two ways of estimating 3.12: the generalized method of moments (GMM) and the control function approach.

Estimation by GMM

Starting from 3.12, we can set the error function, as done by Greene (2003),

$$\mu_i \equiv \mu(y_i, \mathbf{x}_i, y_{2i}, \beta_1, \beta_2) = \frac{y_i}{\exp\left(\mathbf{x}_i'\beta_1 + y_{2i}'\beta_2\right)} - 1.$$

Given instrument z, the population-moment conditions are $\mathbb{E}(\tilde{z}_i\mu(y_i, \mathbf{x}_i, y_{2i}, \beta_1, \beta_2)) = \mathbf{0}$, where $\tilde{z}_i = [\mathbf{x}_i', z_i']$. The value of $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2]$ are the values of β that minimize the quadratic function

$$Q = \left(N^{-1}\sum_i^N z_i'\mu_i\right)' \mathbf{W} \left(N^{-1}\sum_i^N z_i'\mu_i\right),$$

where $\mathbf{W}$ is the weight matrix.

Estimation with Control Function

The main idea behind a control function is to add an error in the model that will allow for overdispersion and endogeneity, as explained by Cameron and Trivedi (2013), which our presentation follows. Again, starting from 3.12, we consider the recursive model, with a new parametrization for y_{2i}:

$$\mathbb{E}(y_i|\mathbf{x}_i, y_{2i}, \epsilon_{1i}) = \exp\left(\mathbf{x}_i'\beta_1 + \beta_2 y_{2i} + \epsilon_{1i}\right)$$
$$y_{2i} = \tilde{\mathbf{z}}_i'\tilde{\beta}_2 + \epsilon_{2i},$$

with $\tilde{\mathbf{z}}_i' = [\mathbf{x}_i' z_i']$. Therefore, if $\mathrm{Corr}(\epsilon_{1i}, \epsilon_{2i}) \neq 0$, the Poisson regression will yield to inconsistent estimates. However, if we assume that ϵ_{1i} is a function of ϵ_{2i}, for instance $\epsilon_{1i} = \alpha\epsilon_{2i} + \omega_i$, with $\omega_i \perp\!\!\!\perp \epsilon_{2i}$, it implies that

$$\mathbb{E}(y_i|\mathbf{x}_i, y_{2i}, \epsilon_{2i}) = \exp\left(\mathbf{x}_i'\beta_1 + \beta_2 y_{21} + \alpha\epsilon_{2i}\right).$$

This setting suggests the following method:

1. We perform an OLS of y_2 on $\tilde{\mathbf{z}}$, and we save the residuals, ϵ_2;

2. We perform a Poisson regression of y_1 on $\mathbf{x}$, y_2 and ϵ_2. This will lead to consistent IV-estimates.

We can finally note that the endogeneity of y_2 can be tested based on $\hat{\alpha}$:

$$\text{H0} : \alpha = 0 \text{ (no endogeneity)}$$
$$\text{H1} : \alpha \neq 0 \text{ (endogeneity).}$$

Chapter 4

Data and Descriptive Statistics

Learning comes from asking "Why do things work like that? Why not some other way?" The world is a very puzzling place. If you're not willing to be puzzled, you just become a replica of someone else's mind.
— Noam Chomsky, *Animating Noam Chomsky*

In this chapter, we will discuss the data that was used to answer our research questions. Since our goal is to explain the influence of economic conditions and marriage market on recidivism, this section will be separated into two parts. First, we will discuss the source of the data on Quebec prison population and highlight some limitations about the dataset. We will present an exploratory analysis by displaying some descriptive statistics. Second, we will provide summary statistics about the variables used to measure the economic conditions and the marriage market.

4.1 Data on Prison Population

4.1.1 Source of the Data

The data was provided by the ministère de la Sécurité publique and consisted of several tables with a unique inmate id as the key between the tables. The covered period starts in 2006 and ends in 2017 - however, we did not have access to the full information for 2006 and 2017 as observations started and ended in the middle of these two years. Still, we used the available information to create our panel dataset (detailed in Section 4.1.2) and removed the observations corresponding to 2006 and 2017 afterwards. The number of admissions recorded each year is shown on Figure 4.1.

Overall, there are 460,928 admissions recorded over the period[1]. Prior to building our panel

[1]Several cases have to be considered. As an admission corresponds to a unique passage in jail, even if the

dataset, we excluded observations with obvious mistakes.

- 21,858 passages were marked as still in custody when it was unlikely, since either the individual had reoffended after or had served for more than two years. These observations were excluded as they seemed due to randomness.
- 73 observations with wrongly inputted regional codes were excluded.
- 3 passages were spent in non-existent detention centers and were removed.

Figure 4.1 – Number of Admissions Over 2006-2017 in Quebec Prisons

In the next section, we will detail the procedure to join the different tables in a way to remove the passages that were not the result of an offense.

4.1.2 Data Management

Join between Tables

For our analysis, we had to connect each passage of every individual and the crime that led to the incarceration, pieces of information available in two different tables. We made the following assumptions:

1. *Many offenses lead to one incarceration*[2];

passage only lasts a few hours, not every passage is the result of a condemnation. More details will be provided in the next section on this matter.

[2]A crime is generally linked with several articles in the Criminal Code.

2. *An individual is incarcerated for the crime for which the difference between the date of trial and the date of admission is the closest.*

Let's consider the following example, given in Table 4.1. First, we merged the observations that have the same trial date.

Table 4.1 – Merging Trial Information

id	offense	trial	region
A	CC0145	2008-01-30	8
A	CC0733	2008-01-30	8
A	ST0004	2010-03-04	8

→

id	offense	trial	region
A	CC0145, CC0733	2008-01-30	8
A	ST0004	2010-03-04	8

Second, suppose we have the following admission information:

Table 4.2 – Admission Information

id	date1	date2
A	2008-02-24	2009-01-01

We then use the SQL inner join function to create the following dataset, in which we compute the difference between the dates. We keep the minimal difference, in absolute value.

Table 4.3 – Join between Trial and Admission

id	date1	date2	offense	trial	region	diff
A	2008-02-24	2009-01-01	CC0145, CC0733	2008-01-30	8	25 days
A	2008-02-24	2009-01-01	ST0004	2010-03-04	8	739 days

Panel Data

We expanded the data to create a panel over the 2007-2016 period. In this part, we will explain how we manipulated the data as well as the underlying assumptions we have made. At first, let's see an example of the data in their original form.

Table 4.4 – Data in Original Form

id	year1	year2	region
A	2008	2008	8
A	2008	2009	8
A	2012	2013	14

To build the panel dataset, we make the following assumptions:

1. *At the time of the offense, the individual lives in the municipality of the trial court*;
2. *Between two prison records, the individual lives where the next offense will be committed*;
3. *After the last sentence, the individual returns where the last crime was committed.*

Since our variables measuring economic conditions are treated annually, we consider only one passage for two or more offenses committed the same year. That is, we keep the information of the last sentence only. The variable *freq* will stand for the number of admissions during the year.

Table 4.5 – Data in Panel Format

panel	id	year1	year2	total	freq	possibility of entering?	prison?	region
2007	A				0	yes	no	8
2008	A	2008	2009	2	2	yes	yes	8
2009	A				0	yes	no	14
2010	A				0	yes	no	14
2011	A				0	yes	no	14
2012	A	2012	2013	2	1	yes	yes	14
2013	A				0	yes	no	14
2014	A				0	yes	no	14
2015	A				0	yes	no	14
2016	A				0	yes	no	14

The same procedure is applied to the civil status and to the number of dependants, where we did a linear approximation between admissions to allow for the number of dependants to vary. Suppose we have information about the number of dependants (n_i and n_j) for respectively $year_i$ and $year_j$, we used the following formula to compute n_k for $k \in (i, j)$:

$$n_k = \left\lfloor \frac{n_i(year_j - year_k) + n_j(year_k - year_i)}{year_j - year_i} \right\rfloor,$$

where $\lfloor \bullet \rfloor$ is the floor function.

If we had had a dataset such as Table 4.6,

Table 4.6 – Data in Original Form: A Longer Sentence

id	year1	year2	region
A	2008	2010	8

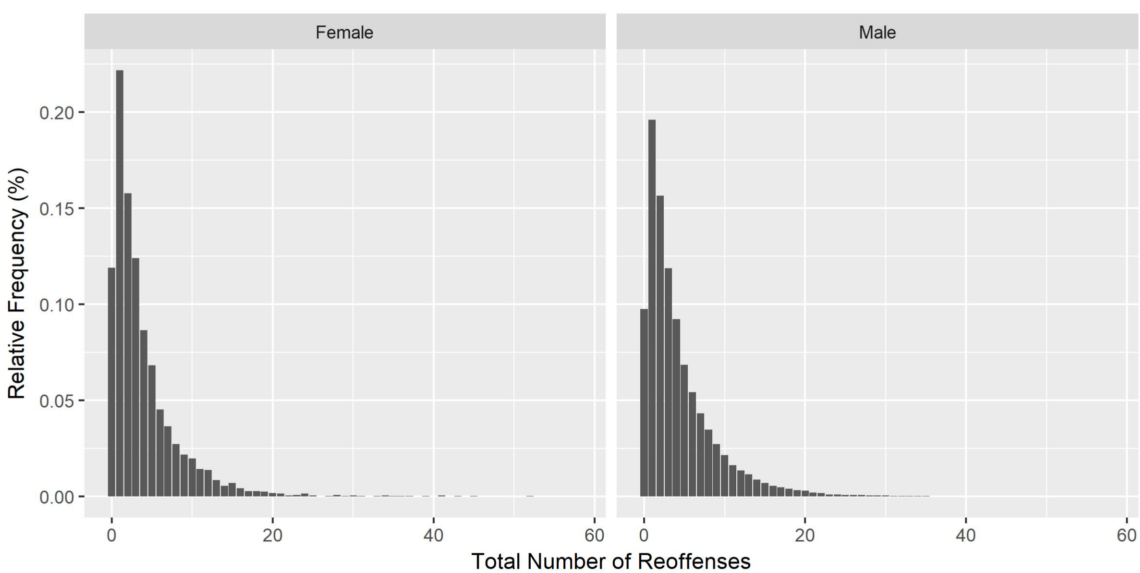
we would not have considered the year 2009 since the individual cannot return to prison for he is already serving a sentence. Such occurrences do not appear in our data. Finally, to complete our recidivism model, we select observations corresponding to reoffenses only[3]. That is, for example, in Table 4.5, we remove the observation with *panel* = 2007. In the end, we were able to track the history of 45,948 men and 4,940 women. After removing the first infraction as we detailed above, our panel contains information about 44,696 men and 4,812 women.

4.1.3 Inmates' Social Characteristics

In this section, we provide descriptive statistics about the offenders' social characteristics. We pay particular attention to our variables of interest, marital status and number of dependants, but also present exclusive exploratory statistics that could potentially be used for future research projects.

As our aim is to model recidivism, both for the probability that it happens and for the average of occurrences over a one-year period, we had to have significant variations among individuals, which is the case. As displayed in Figure 4.2, around 12% of women and 10% of men only visited prison once and around 25% of women and 20% of men reoffended only once. The recorded maximum of recidivism over the 10-year period is 52 for women and 58 for men.

Figure 4.2 – Number of Reoffenses by Sex Over 2006-2017

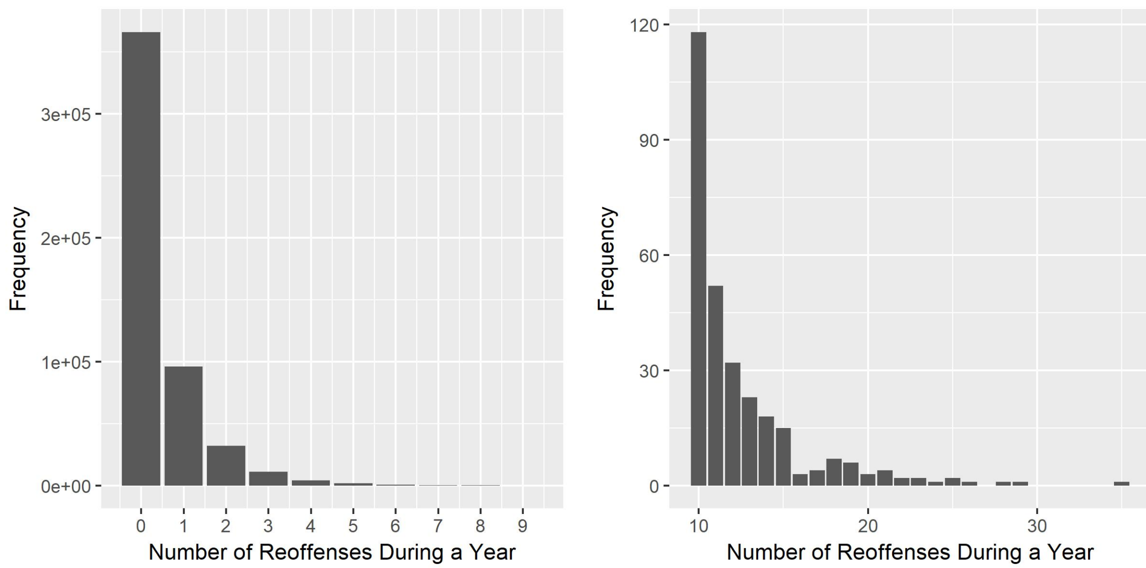


Count models such as Poisson are efficient if there is sufficient variations among the occurrences of an event during a precise period of time. In this thesis, since our economic conditions variables are measured annually, we record the number of offenses over every one-year period.

[3] As an exception, if the first offence is followed by one or more reoffenses during the same year, we keep this observation.

In the dataset under the panel form, as expected, we observe many zeros (recall Table 4.5 as an example[4]). The maximum number of occurrences recorded for a one-year period is 35. Figure 4.3 summarizes this data and was separated for reading purposes.

Figure 4.3 – Number of Reoffenses During a Year Over 2007-2017

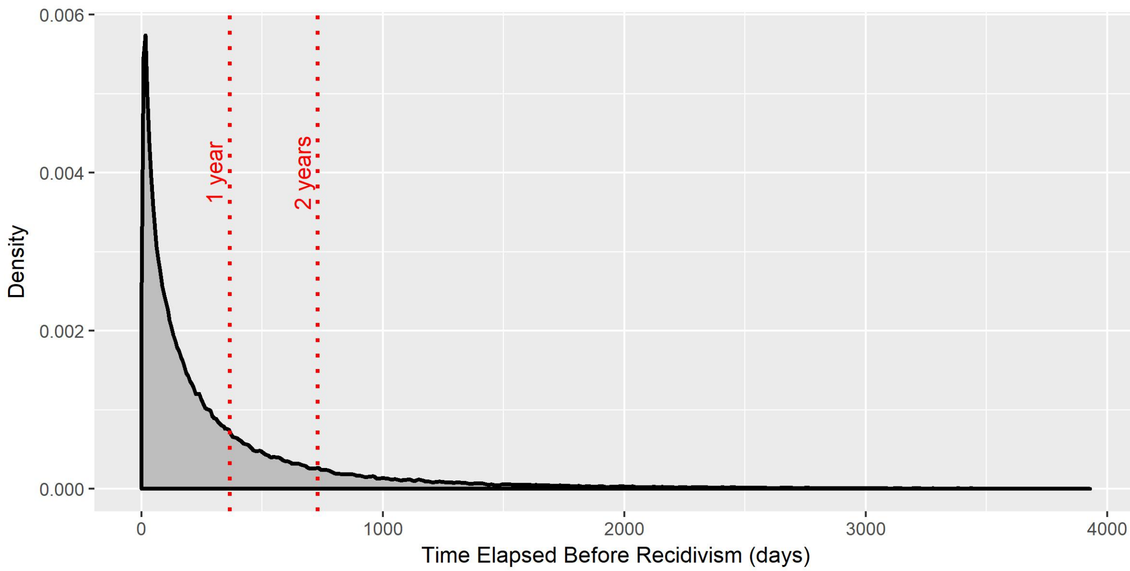


(a) For Number of Reoffenses < 10 (b) For Number of Reoffenses ≥ 10

We may recall that, in this thesis, as first explained in Chapter 1, we consider recidivism to be a return to custody, without regards to the delay between crimes. On the other hand, it is not uncommon for the scientific literature to delineate recidivism over an interval of two years only. Our paper can still be compared to already published studies since, as shown in Figure 4.4, a large proportion of reoffenses are committed over a two-year horizon. In fact, among men recidivists, respectively 72% and 87% of admissions are before one and two years following a prior release.

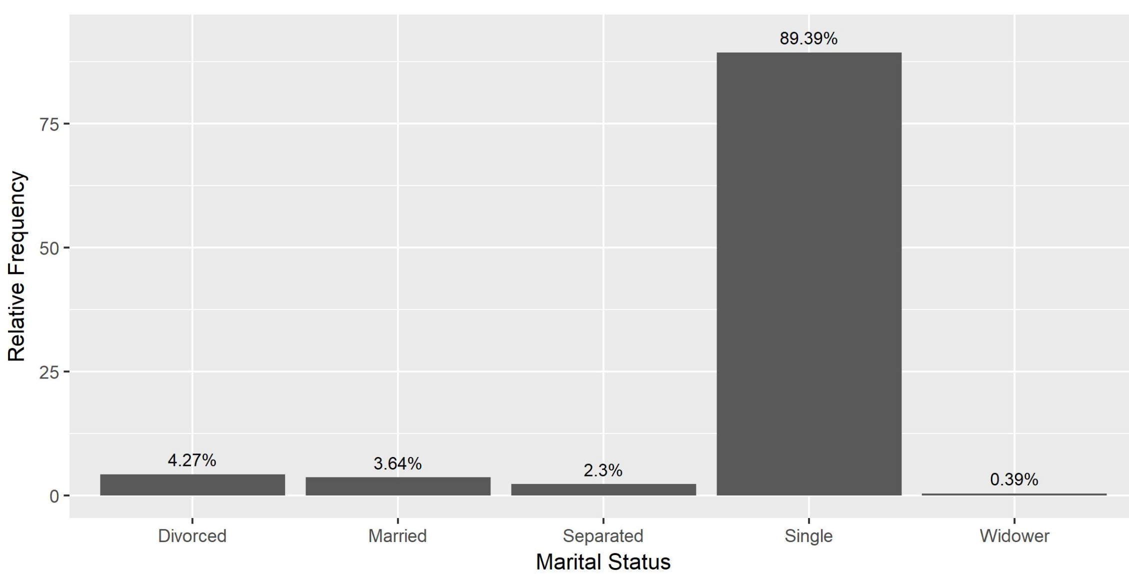
[4] Here, the variable *freq* is plotted.

Figure 4.4 – Time Elapsed before Recidivism among Men Reoffenders



A small sample of our population is married, as exhibits Figure 4.5. Among all the inmates that were admitted between 2006 and the beginning of 2017, 3.64% of the individuals were married, which represents 251 different men in our panel. In the panel dataset, 376,357 observations are individuals without a family, whereas 82,985 observations are individuals with at least one dependant. Once we remove the first offenses, our sample contains 52,945 observations from individuals with families.

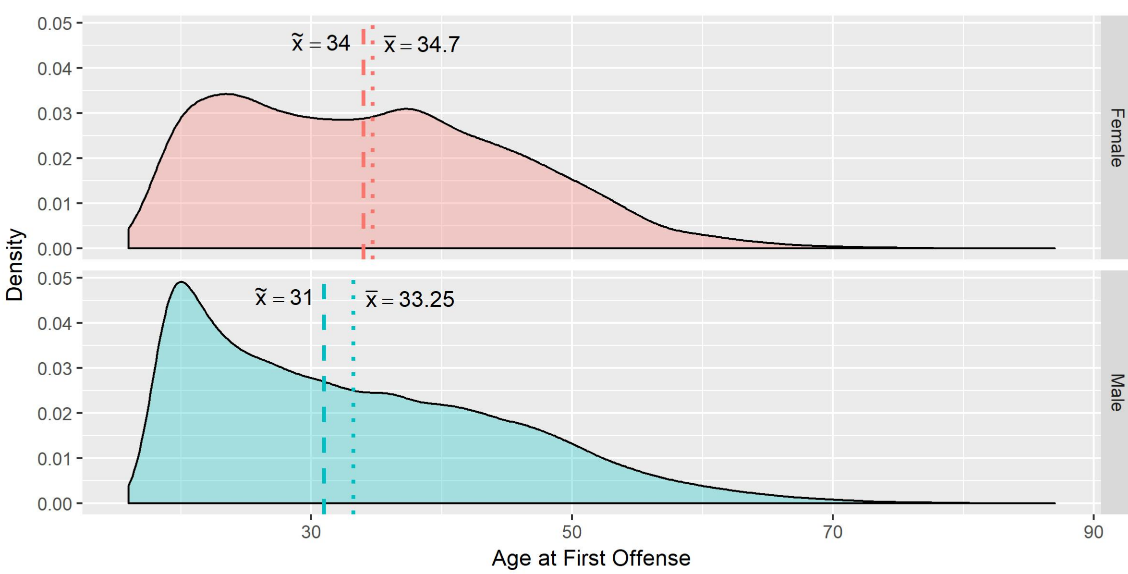
Figure 4.5 – Relative Frequency of Marital Status Among Inmates



Due to the larger proportion of men among inmates, in this thesis, we decided to focus specifically on the men population. Furthermore, men inmates seem to indulge in different behaviours from women, which could have biased our analysis if they were treated together. As an example, we display in Figures 4.6 and 4.7, the density of the age at the time of the first offense by men and women, where we note that the skewness and kurtosis of both distributions are different. The intuition behind the fact that men tend to get involved in crimes sooner than women could be that women generally have to deal with familial duties sooner in their life.

Figure 4.6 – Relative Frequency of Age at First Offense by Sex

Figure 4.7 – Mean and Density of Age at First Offense by Sex
($\tilde{x}$: median; $\bar{x}$: mean)



4.2 Data on Economic Conditions

4.2.1 Unemployment Rate

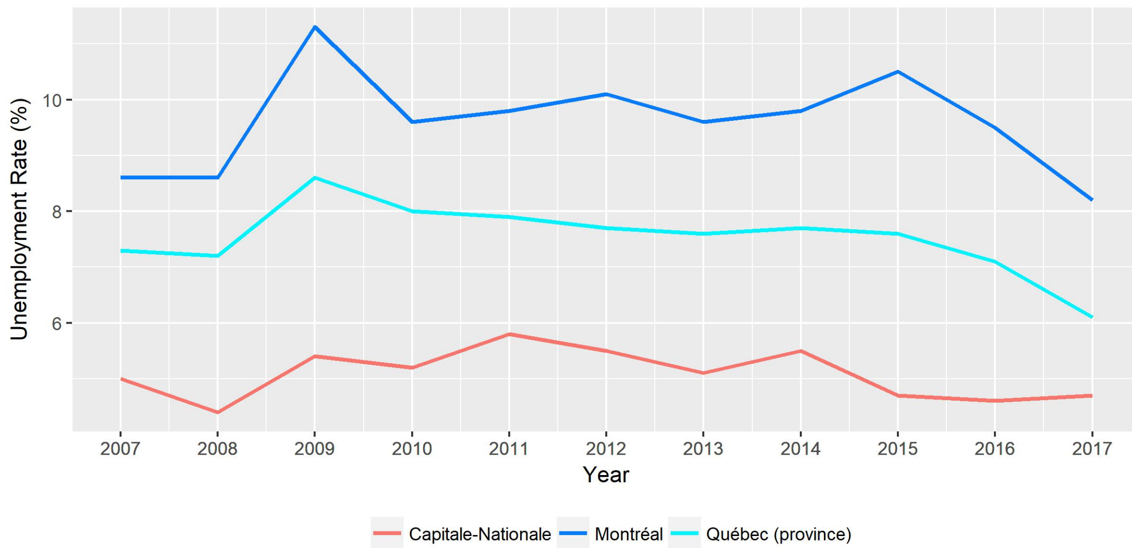
Data on the regional unemployment rate, first collected by Statistics Canada, was found on the website of Institut de la statistique du Québec. We present, in Table 4.7, the means and the standard errors of the unemployment rate from all the regions of Quebec from 2007 to 2017. Note that Côte-Nord and Nord-du-Québec are considered together in the data, which does not induce a bias in our analysis since these two regions would have presented similar numbers, even treated separately.

Table 4.7 – Means and Standard Errors of Regional Unemployment Rates from 2007 to 2017

Code	Region	Mean	Standard Error
1	Bas-Saint-Laurent	8.59	1.06
2	Saguenay–Lac-Saint-Jean	8.58	0.89
3	Capitale-Nationale	5.08	0.44
4	Mauricie	8.51	1.15
5	Estrie	6.78	0.95
6	Montréal	9.60	0.89
7	Outaouais	6.76	0.73
8	Abitibi-Témiscamingue	7.49	1.28
9-10	Côte-Nord and Nord-du-Québec	8.53	1.40
11	Gaspésie–Îles-de-la-Madeleine	14.96	1.90
12	Chaudière-Appalaches	5.20	0.71
13	Laval	7.10	0.86
14	Lanaudière	7.48	0.86
15	Laurentides	7.12	0.89
16	Montérégie	6.47	0.89
17	Centre-du-Québec	6.89	0.81
	Québec (province)	7.53	0.63

In our research, in addition to taking into account the region where an individual is released, we will consider the year of release and the variations in unemployment rates. We display, in Figure 4.8, the unemployment rates from 2007 to 2017 of Capitale-Nationale and Montréal, the two regions with the largest prison population. Moreover, we show the unemployment rate from the Province of Quebec, which is the mean of the unemployment rates from all the regions listed in Table 4.7.

Figure 4.8 – Unemployment Rates from 2007 to 2017



4.2.2 Minimum Wage and Disposable Income

To enrich our measure of economic conditions, we also consider the variations in the minimum wage over the last ten years. The data was found on the website of Commision des normes, de l'équité, de la santé et de la sécurité du travail. To simplify the join between the data frames, we treat the data annually and took the mean of the fluctuations over every one-year period.

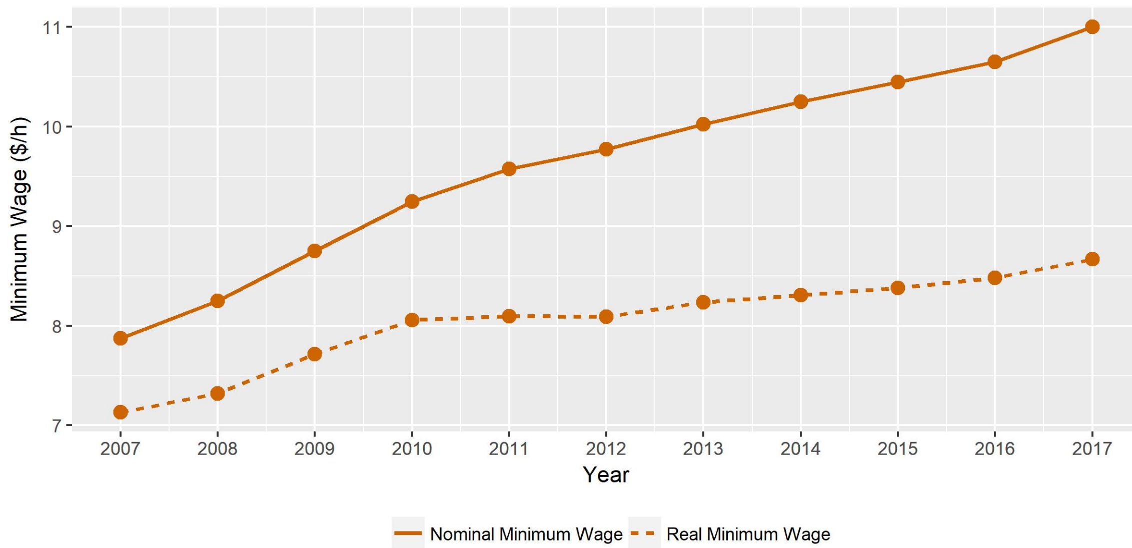
To correct for inflation, we used the annual Consumer Price Index (CPI) for the province of Quebec[5]. The data for CPI was provided by Institut de la statistique du Québec. More precisely, we computed, for year i:

$$\text{Real Wage}_i = 100 \left(\frac{\text{Nominal Wage}_i}{CPI_i} \right). \tag{4.1}$$

We present in Figure 4.9 the history of Quebec's minimum wage.

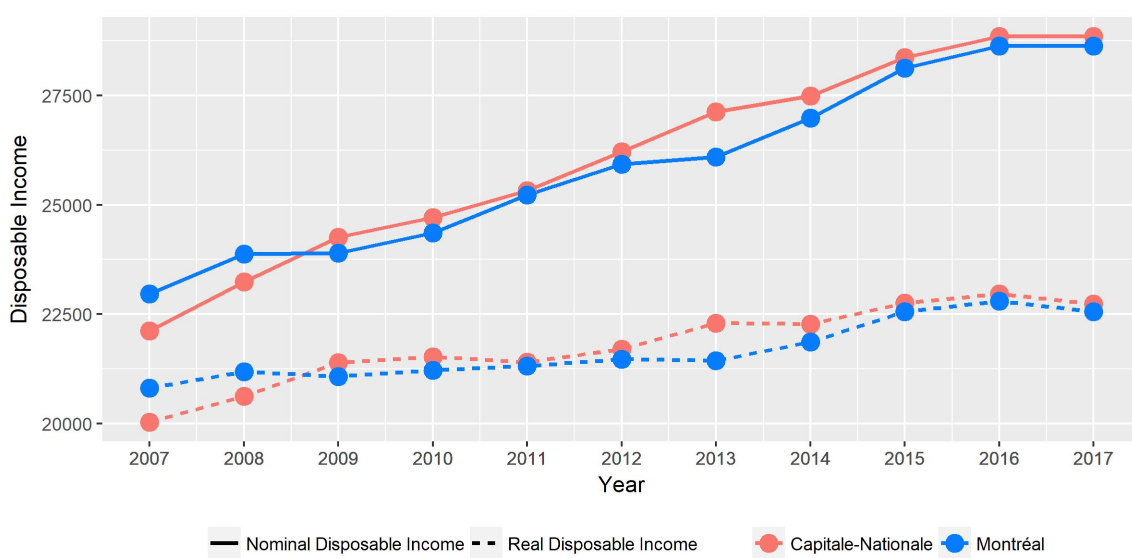
[5]They consider 2002 as the baseline year.

Figure 4.9 – History of the Minimum Wage in the Province of Quebec



To take into account the wealth of the regions, we incorporate in our model the average disposable income of each region over the period 2007-2017, found on the website of Institut de la statistique du Québec. Again, we correct for inflation by performing a computation similar to 4.1. Figure 4.10 shows a preview of the data.

Figure 4.10 – Fluctuations in Disposable Income in Capitale-Nationale and in Montréal from 2007 to 2017



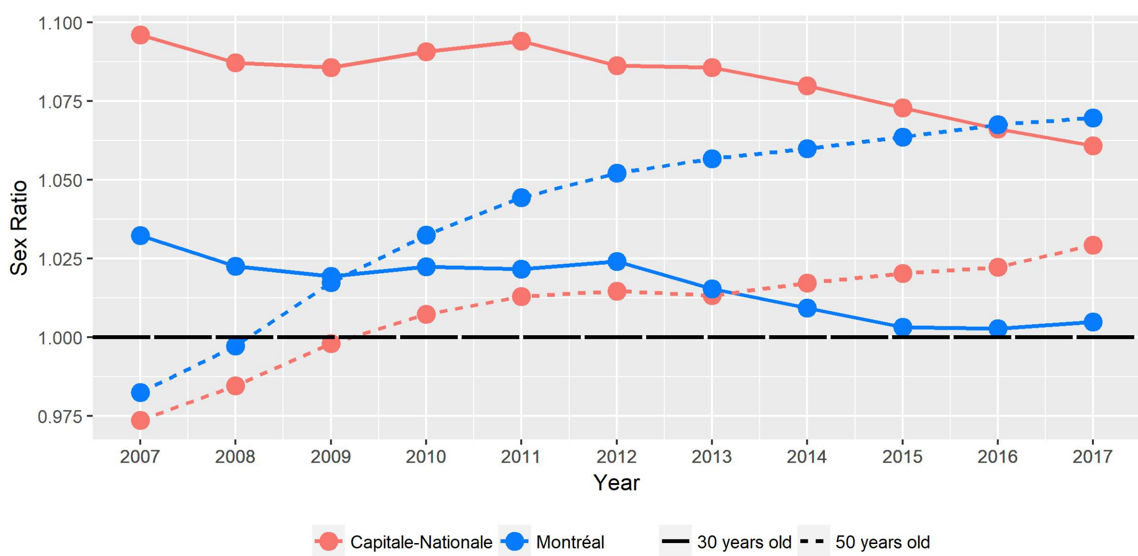
4.2.3 Sex Ratio

The regional sex ratios were calculated from a recent publication from Institut de la statistique du Québec. We adapt the data to the age of the released individual in addition to his geographical location. For example, for an i-year-old individual released in region j in year k, we used the next formula:

$$SR_{ijk} = \frac{\text{Number of men aged from } i-5 \text{ to } i+5 \text{ in region } j \text{ in year } k}{\text{Number of women aged from } i-5 \text{ to } i+5 \text{ in region } j \text{ in year } k}.$$

That is, we supposed a 10-year window on age to evaluate the marriage market for an individual. Among all the possible sex ratios that we could have displayed, we show, in Figure 4.11, the regional sex ratios for individuals aged 30 and 50, in either Capitale-Nationale and Montréal, from 2007 to 2017. Note that the y-axis does not show 0; sex ratios do not vary as much as the graph makes them look to be.

Figure 4.11 – Fluctuations in Sex Ratios for Individuals Aged 30 and 50



Chapter 5

Results

Ne pas se moquer, ne pas se lamenter, ne pas détester, mais comprendre.
— Baruch Spinoza (translated by Frédéric Lenoir)

5.1 Probability of Recidivism

The results from the probit regressions are reported in Tables 5.1 (with endogenous variable *marriage*) and 5.2 (with endogenous variable *family*), and will be analyzed separately. We first perform a naive probit regression in which *marriage* is considered exogenous (model (1)).

We find that being married and the minimum wage are negatively correlated with the probability of recidivism. Indeed, being married decreases the probability of recidivism of 3.70% and a one unit change in the minimum wage decreases this probability of 29.13%. Similarly, one unit change in the unemployment rate diminishes a reoffense probability by 1.25%, which may seem counter-intuitive, as unemployment is generally positively correlated with recidivism. We have to recall that our unemployment measure is regional-based only and does not take into account the age of the individual, his sex or his sector of employment. Moreover, regional unemployment rates might not capture the employment opportunities of the criminal population (Schnepel (2013)). Nevertheless, the negative correlation could be explained by the fact that a region with a higher unemployment rate is generally less wealthy and could lead to fewer opportunities in crime, which is especially true in the range of crimes studied here. The same discussion applies to disposable income, where we find that an increase of 1000$ in the average regional disposable income is linked with a 3.31% rise in the likelihood of recidivism.

The effect of age on the probability of recidivism is intuitive and significant. We find that moving from one age class to another decreases the likelihood of reoffending by between 9.44% and 21.13%. We also consider regional fixed-effects in this regression. Being released in

Capitale-Nationale, Chaudières-Appalaches, Côte-Nord or Montérégie decreases significantly the probability of recidivism (compared to being released in Abitibi-Témiscamingue), whereas the opposite is found for Estrie, Gaspésie-Îles-de-la-Madeleine, Mauricie and Montréal. No significant effect is found for the other regions.

Before computing models (2) and (3) in which we use the sex ratio as an instrument for *marriage*, we perform a weak identification test for our instrument. The Cragg-Donald Wald F-Statistic is 74.441 and, therefore, sex ratio seem to be a good instrument for *marriage*. Similarly, we perform a Wald test for the endogeneity of *marriage*, whose null hypothesis is that *marriage* is exogenous with the probability of recidivism. With the p-values reported in Table 5.1, our data fails to reject the null hypothesis for both (2) and (3). This is probably explained by the fact that we have rather few observations with married individuals. Altogether, the Wald test does not allow us to determine if *marriage* is endogenous.

Model (2), a probit regression with an instrumental variable, deal with *marriage* as if it was continuous. Except from being married, the effect of economic conditions are amplified compared to the effects found with the naive probit. Again, since *marriage* is binary, the estimated coefficients are probably biased. Model (3), the bivariate probit, is plausibly the most adequate model. The unemployment rate, the minimum wage and the disposable income have rather the same effect as in the naive probit, as well as for the age and regional fixed-effects. The interesting fact lies in *marriage* which, once instrumented, is found not to be significant. With the corresponding standard error, we can build the confidence interval for the marginal effect of *marriage*, that we denote $\tilde{\beta}_{\text{marriage}}$:

$$\mathbb{P}\left(\tilde{\beta}_{\text{marriage}} \in [-0.181, 0.065]\right) = 95\%,$$

which is a rather wide interval. Therefore, we cannot conclude whether or not marriage has a significant impact, reasonably because of the lack of observations of married individuals. The marginal effects of the economic conditions are consequent with the discussion above. A one unit change in the unemployment rate decreases the probability of recidivism by 1.25%, a one unit change in the minimum wage decreases the probability of recidivism by 29.13% and an increase of 1000$ in the average regional disposable income increases the probability by 3.31%.

Table 5.1 – Results from Probit Regressions (Endogenous Variable: *marriage*)

		(1) Naive Probit			(2) Probit with IV			(3) Bivariate Probit		
Dependent Variable: Recidivism		Coef	SE	mfx	Coef	SE	mfx	Coef	SE	mfx
Economic Conditions										
	marriage	−0.1023***	0.0175	−0.0370	−1.8692	1.1405	−1.8692	−0.1610	0.1735	−0.0582
	unemployment rate	−0.0347***	0.0041	−0.0125	−0.0318***	0.0049	−0.0318	−0.0346***	0.0041	−0.0125
	minimum wage	−0.8060***	0.0178	−0.2913	−0.7715***	0.0510	−0.7715	−0.8061***	0.0178	−0.2913
	disposable income (k$)	0.0917***	0.0098	0.0331	0.0836***	0.0122	0.0836	0.0915***	0.0098	0.0331
Age [18-25]										
	[26 - 35]	−0.2480***	0.0083	−0.0944	−0.2128***	0.0320	−0.2128	−0.2472***	0.0086	−0.0940
	[36 - 45]	−0.2984***	0.0092	−0.1128	−0.2198***	0.0615	−0.2198	−0.2962***	0.0113	−0.1119
	[46 - 55]	−0.4022***	0.0102	−0.1496	−0.2990***	0.0809	−0.2990	−0.3994***	0.0133	−0.1485
	[55 -]	−0.5899***	0.0146	−0.2113	−0.4052***	0.1390	−0.4052	−0.5847***	0.0215	−0.2095
Region (Abitibi-Témiscamingue)										
	Bas-Saint-Laurent	0.0429	0.0327	0.0156	−0.1162	0.1089	−0.1162	0.0378	0.0359	0.0138
	Capitale-Nationale	−0.2608***	0.0196	−0.0896	−0.3866***	0.0761	−0.3866	−0.2653***	0.0237	−0.0912
	Centre-du-Québec	0.0357	0.0421	0.0130	−0.1234	0.1120	−0.1234	0.0306	0.0447	0.0111
	Chaudière-Appalaches	−0.1313***	0.0287	−0.0464	−0.2613***	0.0850	−0.2613	−0.1357***	0.0315	−0.0480
	Côte-Nord	−0.0465*	0.0282	−0.0167	−0.1803**	0.0898	−0.1803	−0.0509	0.0311	−0.0183
	Estrie	0.1490***	0.0264	0.0551	−0.0041	0.1069	−0.0041	0.1442***	0.0298	0.0533
	Gaspésie–Îles-de-la-Madeleine	0.2589***	0.0419	0.0967	0.0983	0.1196	0.0983	0.2540***	0.0442	0.0950
	Lanaudière	0.0153	0.0226	0.0056	−0.1061	0.0825	−0.1061	0.0114	0.0254	0.0041
	Laurentides	−0.0291	0.0190	−0.0105	−0.1476*	0.0786	−0.1476	−0.0330	0.0223	−0.0119
	Laval	0.0333	0.0226	0.0121	−0.0670	0.0706	−0.0670	0.0300	0.0246	0.0109
	Mauricie	0.1667***	0.0276	0.0617	0.0039	0.1137	0.0039	0.1616***	0.0312	0.0599
	Montréal	0.0404*	0.0210	0.0147	−0.0836	0.0845	−0.0836	0.0365	0.0241	0.0133
	Montérégie	−0.1591***	0.0175	−0.0559	−0.2709***	0.0701	−0.2709	−0.1630***	0.0210	−0.0573
	Nord-du-Québec	0.1613	0.1245	0.0597	0.1855	0.1479	0.1855	0.1624	0.1248	0.0602
	Outaouais	0.0313	0.0212	0.0114	−0.0969	0.0867	−0.0969	0.0272	0.0245	0.0099
	Saguenay–Lac-Saint-Jean	0.0887***	0.0242	0.0325	−0.0499	0.0960	−0.0499	0.0843***	0.0274	0.0310
Constant		4.8424***	0.1246		4.8410***	0.2004		4.8496***	0.1261	
Legend	*** 1% ** 5% * 10%									
Observations			284,956			284,956			284,956	
Wald Test					2.1000	(p-value: 0.1477)		0.1158	(p-value: 0.7336)	
Notes:	*marriage* is instrumented by sex ratios in (2) and (3). Standard errors (SE) are adjusted for clusters. mfx are the average marginal effects.									

We perform the same regressions, this time using *family* instead of *marriage*. The results are reported in Table 5.2. In model (1), the naive probit, *family* is found to be positively correlated with recidivism: having a family increases the likelihood of recidivism by 1.73%. The effects from the other economic conditions, age and regional fixed-effects have pretty much the same magnitude as previously.

In models (2) and (3), *family* is instrumented with sex ratios. The Cragg-Donald Wald F-Statistic for sex ratio is 47.537 and therefore can be considered as a good instrument. We proceed to a Wald test to determine whether *family* is endogenous. Model (2) fails to reject the null hypothesis, whereas, in the bivariate probit, the p-value of 0.0001 strongly advocates against exogeneity[1]. This leads us to invalidate the estimated coefficients from model (1) as this model does not take into account the endogeneity of *family*.

The estimations provided by model (2) should be taken with a grain of salt since $family$ is a binary predictor. Rather, model (2) allows us to verify the robustness of the estimations from model (3), which is more appropriate when the endogenous predictor is binary.

Once instrumented, having a family decreases the probability of recidivism of 18.26%, which is consistent with Becker's economic theory, since having a family increases the *cost* of a crime. We can therefore conclude with certainty that the effect of having a family has a negative causal impact on the probability of reoffending. In this particular setting, the economic variables and age have a similar impact as when *marriage* was used instead of *family*.

The fact of being released in Bas-Saint-Laurent, Capitale-Nationale, Centre-du-Québec, Chaudière-Appalaches, Côte-Nord, Lanaudière, Laurentides, Laval, Montréal and Montérégie is linked with a smaller probability of recidivism compared to being released in Abitibi-Témiscamingue. At the opposite, being released in Gaspésie-Îles-de-la-Madeleine increases the probability of reoffense by 4.84%.

[1]This probably comes from the fact that a significant part of our observations are about individuals who have a family, whereas our sample of married individual was small.

Table 5.2 – Results from Probit Regressions (Endogenous Variable: *family*)

		(1) Naive Probit			(2) Probit with IV			(3) Bivariate Probit		
Dependent Variable: Recidivism		Coef	SE	mfx	Coef	SE	mfx	Coef	SE	mfx
Economic Conditions										
	family	0.0479***	0.0080	0.0173	1.0223*	0.5926	1.0223	−0.5095***	0.1322	−0.1826
	unemployment rate	−0.0349***	0.0041	−0.0126	−0.0354***	0.0041	−0.0354	−0.0324***	0.0041	−0.0116
	minimum wage	−0.8055***	0.0179	−0.2911	−0.7463***	0.0760	−0.7463	−0.7877***	0.0195	−0.2822
	disposable income (k$)	0.0917***	0.0099	0.0331	0.0847***	0.0129	0.0847	0.0899***	0.0095	0.0322
Age [18-25]										
	[26 - 35]	−0.2548***	0.0083	−0.0971	−0.3460***	0.0442	−0.3460	−0.1854***	0.0201	−0.0692
	[36 - 45]	−0.3112***	0.0093	−0.1177	−0.4750***	0.0855	−0.4750	−0.1947***	0.0317	−0.0726
	[46 - 55]	−0.4139***	0.0102	−0.1540	−0.5278***	0.0508	−0.5278	−0.3197***	0.0276	−0.1170
	[55 -]	−0.6020***	0.0146	−0.2157	−0.6285***	0.0191	−0.6285	−0.5487***	0.0220	−0.1919
Region (Abitibi-Témiscamingue)										
	Bas-Saint-Laurent	0.0632*	0.0328	0.0230	0.3044**	0.1488	0.3044	−0.0813*	0.0487	−0.0297
	Capitale-Nationale	−0.2406***	0.0196	−0.0825	0.0301	0.1781	0.0301	−0.3832***	0.0391	−0.1328
	Centre-du-Québec	0.0563	0.0422	0.0205	0.3065*	0.1565	0.3065	−0.0933*	0.0561	−0.0340
	Chaudière-Appalaches	−0.1123***	0.0288	−0.0396	0.1337	0.1588	0.1337	−0.2486***	0.0438	−0.0884
	Côte-Nord	−0.0294	0.0282	−0.0105	0.1657	0.1244	0.1657	−0.1405***	0.0391	−0.0508
	Estrie	0.1681***	0.0266	0.0620	0.3840***	0.1275	0.3840	0.0318	0.0433	0.0117
	Gaspésie–Îles-de-la-Madeleine	0.2791***	0.0419	0.1041	0.5059***	0.1333	0.5059	0.1301**	0.0567	0.0484
	Lanaudière	0.0318	0.0227	0.0115	0.2314*	0.1231	0.2314	−0.0859**	0.0371	−0.0313
	Laurentides	−0.0119	0.0190	−0.0043	0.2010	0.1321	0.2010	−0.1349***	0.0357	−0.0489
	Laval	0.0481**	0.0226	0.0175	0.2371**	0.1160	0.2371	−0.0645*	0.0361	−0.0236
	Mauricie	0.1869***	0.0277	0.0691	0.4154***	0.1345	0.4154	0.0418	0.0459	0.0155
	Montréal	0.0576***	0.0210	0.0209	0.2616**	0.1245	0.2616	−0.0643*	0.0368	−0.0235
	Montérégie	−0.1429***	0.0175	−0.0501	0.0619	0.1332	0.0619	−0.2524***	0.0314	−0.0897
	Nord-du-Québec	0.1570	0.1247	0.0578	0.1215	0.1428	0.1215	0.1682	0.1214	0.0628
	Outaouais	0.0464**	0.0213	0.0168	0.2102**	0.1009	0.2102	−0.0513	0.0322	−0.0188
	Saguenay–Lac-Saint-Jean	0.1062***	0.0243	0.0389	0.3072**	0.1207	0.3072	−0.0170	0.0393	−0.0063
Constant		4.8189***	0.1250		4.2202***	0.6074		4.8502	0.1237	
Legend *** 1% ** 5% * 10%										
Observations			284,956			284,956			284,956	
Wald Test					2.2000	(p-value: 0.1381)		15.0324	(p-value: 0.0001)	

Notes: *family* is instrumented by sex ratios in (2) and (3).
Standard errors (SE) are adjusted for clusters.
mfx are the average marginal effects.

5.2 Number of Reoffenses

The results from the count models are reported in Table 5.3. Count models were run using either *marriage* or *family* as a predictor and the dependent variable is the number of offenses committed over every one-year period.

In model (1), in which *marriage* is considered exogenous with the dependent variable, we find that being married is associated with 0.0877 fewer offenses per year, other things equal. In model (2), estimated by GMM and where *marriage* is instrumented by sex ratios, being married is not a significant predictor anymore. We can build a confidence interval around the marginal effect of *marriage*:

$$\mathbb{P}\left(\tilde{\beta}_{\text{marriage}} \in [-0.6382, 2.3148]\right) = 95\%,$$

which is too large an interval for us to be able to conclude. The other economic conditions are robust to the model specifications and have the same order of magnitude in either model (1) or (2). One unit change in the unemployment rate decreases the number of offenses by 0.0283, a 1$ increase in the real minimum wage is associated with 0.5290 fewer offenses while a 1000$ augmentation in the average regional disposable income increases the number of offenses by 0.0209. Again, this may sound counter-intuitive, but we may recall that the average disposable income measures regional wealth, and wealth yields to more opportunities of crime.

The effect of age from models (1) and (2) is comparable in size, even if the average marginal effects computed while *marriage* is instrumented are slightly greater. Compared to the [18−25] age bracket, we find that belonging to a higher age group significantly decreases the number of reoffenses by year by between 0.2436 and 0.5758. Regional fixed-effects are believed to be biased because of the endogeneity of *marriage*: using sex ratio as an instrumental variable, we find that only in Laval and Montréal do individuals tend to reoffend more over one year, whereas more significant coefficients were estimated with the naive approach.

Similarly as for binary regressions, we then consider having a family as a predictor for the number of offenses committed. In model (3), a naive Poisson regression with *family* considered as exogenous, we find that having a family is associated with 0.0355 additionnal offenses per year, while, once instrumented, the effect disappears. We can build the following interval around the marginal effect of having a family:

$$\mathbb{P}\left(\tilde{\beta}_{\text{family}} \in [-0.9799, 1.5636]\right) = 95\%.$$

As this interval is quite wide, we cannot conclude whether or not having a family has an impact on the number of reoffenses, the data not being precise enough.

The economic conditions variables and age groups have substantially the same effect over the number of reoffenses as above. We suspect the coefficients of the regional fixed variables to be biased in model (3) since they are not robust to specification. Indeed, in model (4), only being released in Estrie (compared to Abitibi-Témiscamingue) as a significant impact with, *ceteris paribus*, 0.2054 additional offenses per year. Notice that, either for *marriage* or *family*, most regional effects are not significant once the model is estimated with an instrumental variable. This comes from the lack of variance in the yearly frequency of recidivism among certain regions. We suspect that, for regions with fewer observations like Côte-Nord, Gaspésie-Îles-de-la-Madeleine and Nord-du-Québec, very few individuals have reoffended many times on an annual basis, thus leading to identification problems. To verify this hypothesis, we clustered regions into four groups, where each of these groups had at least 60,000 observations. The results showed that correcting for the lack of variance among regions led to significant coefficients for the regional fixed-effects in models (2) and (4) and, in addition, the coefficients for *marriage*, *family*, age and all the economic conditions were in the same order of magnitude as before.

Table 5.3 – Results from Poisson Regressions

	marriage						family					
	(1) Naive Poisson			(2) Poisson with IV			(3) Naive Poisson			(4) Poisson with IV		
(Method of Estimation)				(GMM)						(Control Function)		
Dependent Variable: Number of Reoffenses per Year	Coef	SE	mfx	Coef	SE	mfx	Coef	SE	mfx	Coef	SE	mfx
Economic Conditions												
marriage	−0.1357***	0.0269	−0.0877	1.2973	1.1657	0.8383						
family							0.0550***	0.0111	0.0355	0.4509	1.0026	0.2918
unemployment rate	−0.0420***	0.0045	−0.0271	−0.0438***	0.0052	−0.0283	−0.0423***	0.0045	−0.0273	−0.0291***	0.0061	−0.0180
minimum wage	−0.8226***	0.0137	−0.5315	−0.8187***	0.0159	−0.5290	−0.8222***	0.0137	−0.5313	−0.8649***	0.0186	−0.5600
disposable income (k$)	0.0274***	0.0076	0.0177	0.0323***	0.0105	0.0209	0.0276***	0.0076	0.0179	0.0715***	0.0110	0.0464
Age [18-25]												
[26 - 35]	−0.2727***	0.0104	−0.2046	−0.3079***	0.0526	−0.2436	−0.2808***	0.0105	−0.2119	−0.3596***	0.1127	−0.2404
[36 - 45]	−0.3381***	0.0118	−0.2459	−0.4308***	0.1327	−0.3217	−0.3535***	0.0120	−0.2577	−0.4681***	0.1906	−0.2883
[46 - 55]	−0.4891***	0.0142	−0.3316	−0.6095***	0.1707	−0.4194	−0.5034***	0.0143	−0.3423	−0.6084***	0.1472	−0.3758
[55 -]	−0.7642***	0.0237	−0.4580	−0.9850***	0.2985	−0.5758	−0.7797***	0.0236	−0.4687	−0.8585***	0.0768	−0.5010
Region (Abitibi-Témiscamingue)												
Bas-Saint-Laurent	−0.1511***	0.0364	−0.0844	0.0447	0.2597	0.0237	−0.1269***	0.0364	−0.0703	0.0431	0.2554	−0.0329
Capitale-Nationale	−0.1711***	0.0244	−0.0946	−0.0033	0.2208	−0.0017	−0.1480***	0.0245	−0.0812	−0.0520	0.2612	−0.0853
Centre-du-Québec	−0.1146**	0.0496	−0.0652	0.0774	0.2568	0.0417	−0.0902*	0.0496	−0.0509	0.0800	0.2661	−0.0153
Chaudière-Appalaches	−0.2431***	0.0343	−0.1299	−0.0756	0.2222	−0.0378	−0.2209***	0.0343	−0.1170	−0.0525	0.2474	−0.0828
Côte-Nord	0.0914***	0.0354	0.0576	0.2574	0.2191	0.1523	0.1110***	0.0354	0.0693	0.1593	0.2021	0.0486
Estrie	0.2000***	0.0291	0.1333	0.3779	0.2339	0.2382	0.2224***	0.0291	0.1470	0.3924*	0.2361	0.2054
Gaspésie–Îles-de-la-Madeleine	0.1113**	0.0519	0.0709	0.3007	0.2539	0.1820	0.1356***	0.0519	0.0857	0.2549	0.2590	0.1002
Lanaudière	0.0096	0.0281	0.0058	0.1584	0.1950	0.0890	0.0291	0.0281	0.0174	0.1445	0.2090	0.0386
Laurentides	0.1466***	0.0241	0.0950	0.2936	0.1922	0.1770	0.1666***	0.0241	0.1070	0.2508	0.2185	0.1059
Laval	0.1503***	0.0282	0.0976	0.2717*	0.1578	0.1619	0.1681***	0.0282	0.1080	0.2594	0.1993	0.1192
Mauricie	0.0957***	0.0294	0.0605	0.2885	0.2549	0.1735	0.1195***	0.0294	0.0749	0.2835	0.2509	0.1190
Montréal	0.2682***	0.0241	0.1852	0.4164**	0.1925	0.2679	0.2882***	0.0242	0.1972	0.3233	0.2148	0.1594
Montérégie	−0.0422*	0.0219	−0.0249	0.1021	0.1879	0.0558	−0.0233	0.0219	−0.0136	0.0412	0.2004	−0.0212
Nord-du-Québec	−0.0064	0.1435	−0.0038	−0.0540	0.1711	−0.0273	−0.0101	0.1435	−0.0059	0.1036	0.1505	0.0723
Outaouais	0.0191	0.0251	0.0116	0.1726	0.1998	0.0977	0.0367	0.0250	0.0221	0.1633	0.1737	0.0610
Saguenay–Lac-Saint-Jean	0.0276	0.0274	0.0169	0.1931	0.2169	0.1105	0.0483*	0.0273	0.0292	0.1640	0.2157	0.0469
Constant	6.2090***	0.1214		5.9264***	0.3983		6.1795***	0.1215		5.4156	0.2988	
Legend *** 1% ** 5% * 10%												
Observations		284,956			284,956			284,956			284,956	

Notes: *mariage* and *family* are instrumented by sex ratios in (2) and (4).
Standard errors (SE) are adjusted for clusters.
mfx are the average marginal effects.

Chapter 6

Conclusion

Although the flow of prisoners in and out of Quebec detention centers each year is significant and many of the offenders are recidivists, no academic study had ever been conducted on this population with a focus on recidivism. Besides, the literature about recidivism mainly centers on American or European prison populations or often studies a specific type of crime, such as sexual delinquency. To find out more about Quebec offenders' predictors of recidivism, the ministère de la Sécurité publique provided us an exclusive dataset with information about all offenders that were sentenced for two years or less from 2006 to 2017. We consider recidivism as the simple fact of returning to jail, with no regards to the crime committed nor the time gap between admissions.

In this thesis, we studied the impact of regional economic conditions on recidivism: unemployment rate, minimum wage and average disposable income. Moreover, we included in our models domestic variables: being married and having a family. Since these two variables are potentially endogenous with recidivism, they were instrumented by regional sex ratios that were adapted to the age of each individual. The intuition behind this idea is that, for a male offender being released in a region with a high sex ratio (more men relatively to women), competition on the marriage market hampers his chances of finding a partner or founding a family, whereas these two situations would deter him from reoffending according to Becker's theory.

Our aim was twofold: to find out whether these conditions have an impact on the probability of recidivism and whether they affect the number of offenses committed over one-year period. First, we used bivariate probit regressions to estimate the probability of recidivism. We were unable to conclude about the impact of being married since the standard error was too large because of the small sample of married individuals in the dataset. On the other hand, we found that having a family reduces the probability of reoffense by 18.26%. We also found that a 1% rise in the regional employment rate and that a 1000$ rise in the average regional income increase the probability of recidivism by respectively 1.20% and 3.27%[1].

Second, we used a Poisson regression with instrumental variables to estimate the number of reoffenses per year. Although we found no effect of being married or having a family, we estimated that a one unit change in the unemployment rate is associated with 0.0232 fewer reoffenses, a raise of 1$ in the minimum wage decreases the number of reoffenses by 0.5445 while an increase of 1000$ in the average regional disposable income leads to 0.0337 less offenses per year[2].

In both models, we found that older individuals were less prone to recidivism, either for the probability of recidivism than for the number of reoffenses committed. We also pointed out similar behaviours within the regions at significant levels. In the light of these findings, correctional policies could be adapted to the individual's personal characteristics. In order to encourage social rehabilitation and to reduce costs of imprisonment, it appears that individuals without a family or living in wealthier regions would require sustained efforts from correctional services.

Like any other study, our research faces limitations due to the data as well as the methodology used. Since we do not have information about the individuals between admissions, we made a series of assumptions to infer about the region they live in, their matrimonial status and their number of dependants between two offenses. Our measure of economic conditions might also be biased: for the unemployment rate, for example, we used an aggregate regional rate that was not directly adapted to the released individual. Furthermore, we ignored left and right censoring in the data. For example, since our observation period started in 2006-2007, it is likely that individuals had gone to prison before it was recorded in our dataset. Also, some individuals might have been lost during the study: some of them might have been sent to a federal penitentiary or may have died, thus preventing them from reappearing in the dataset. As these left-censored individuals were considered *released* in our analysis, a more in-depth study is needed to correct for censoring.

As regards to methodology, in our analysis of being married, we had to deal with a small sample of married individuals, which led us to an ambiguous result. Regional sex ratios were used to instrument the probability of getting married. However, sex ratios are also correlated with the probability of being part of a couple, an unobserved variable in our dataset. That is, it is possible that sex ratios are correlated with the error term in this case. The problem might have been worked around when we dealt with having a family instead. Finally, zero-inflated count models could have been more appropriate since a large part of offenders only reoffend once. Additional work is required to predict the number of occurrences in prison perfectly.

[1] We take the mean of the effects found in both regressions, even if the marginal effects were in the same order of magnitude, see Tables 5.1 and 5.2, models (3).

[2] We take the mean from Table 5.3, models (2) and (4).

We hope that this first econometric research focusing on Quebec male prison population points the way towards other fascinating and unanswered questions. As we briefly described sooner, women's behaviour tends to differ from men's; it would be enlightening to expound if women offenders are affected by economic conditions and marriage market the same way that men are. It would also be worth considering other variables in such a study: inmates' education or prison occupation rates, to mention just a few. In the present study, we did not discriminate by crime nor by crime types. An interesting idea would be to see if, in the case of recidivism, the crime committed is more serious than the others previously committed, and if peer effects are involved in that matter.

Bibliography

Act Respecting the Quebec Correctional System (2018): [Online; accessed 11-May-2018].

Andersen, S. H., L. H. Andersen, and P. E. Skov (2015): "Effect of marriage and spousal criminality on recidivism," *Journal of Marriage and Family*, 77, 496–509.

Andrews, D. A. and J. Bonta (2010): *The psychology of criminal conduct*, Routledge.

Angrist, J. (2002): "How do sex ratios affect marriage and labor markets? Evidence from America's second generation," *The Quarterly Journal of Economics*, 117, 997–1038.

Bales, W. D. and D. P. Mears (2008): "Inmate social ties and the transition to society: Does visitation reduce recidivism?" *Journal of Research in Crime and Delinquency*, 45, 287–321.

Baum, C. F., Y. Dong, A. Lewbel, T. Yang, et al. (2012): "Binary choice models with endogenous regressors," in *Stata Conference*.

Becker, G. S. (1968): "Crime and punishment: An economic approach," in *The economic dimensions of crime*, Springer, 13–68.

Benda, B. B. (2005): "Gender differences in life-course theory of recidivism: A survival analysis," *International Journal of Offender Therapy and Comparative Criminology*, 49, 325–342.

Bonta, J., B. Pang, and S. Wallace-Capretta (1995): "Predictors of recidivism among incarcerated female offenders," *The Prison Journal*, 75, 277–294.

Cameron, A. C. and P. K. Trivedi (2005): *Microeconometrics: methods and applications*, Cambridge university press.

——— (2013): *Regression analysis of count data*, vol. 53, Cambridge university press.

Chalfin, A. and J. McCrary (2017): "Criminal deterrence: A review of the literature," *Journal of Economic Literature*, 55, 5–48.

Chen, M. K. and J. M. Shapiro (2007): "Do harsher prison conditions reduce recidivism? A discontinuity-based approach," *American Law and Economics Review*, 9, 1–29.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002): "Marriage market, divorce legislation, and household labor supply," *Journal of political Economy*, 110, 37–72.

Commision des normes, de l'équité, de la santé et de la sécurité du travail (2018): "Historique du salaire minimum," [Online; accessed 7-May-2018].

Controlled Drugs and Substances Act (S.C. 1996, c. 19) (1996): [Online; accessed 11-May-2018].

Corman, H. and N. Mocan (2005): "Carrots, sticks, and broken windows," *The Journal of Law and Economics*, 48, 235–266.

Criminal Code (R.S.C., 1985, c. C-46) (1985): [Online; accessed 11-May-2018].

Denver, M., G. Siwach, and S. D. Bushway (2017): "A new look at the employment and recidivism relationship through the lens of a criminal background check," *Criminology*, 55, 174–204.

Drago, F., R. Galbiati, and P. Vertova (2009): "The deterrent effects of prison: Evidence from a natural experiment," *Journal of political Economy*, 117, 257–280.

——— (2011): "Prison conditions and recidivism," *American law and economics review*, 13, 103–130.

Edlund, L., H. Li, J. Yi, and J. Zhang (2013): "Sex ratios and crime: Evidence from China," *Review of Economics and Statistics*, 95, 1520–1534.

Fernandez, J., T. Holman, and J. V. Pepper (2014): "The Impact of Living-Wage Ordinances on Urban Crime," *Industrial Relations: A Journal of Economy and Society*, 53, 478–500.

Freeman, R. (1983): "Crime and unemployment," *Crime and public policy*.

Freeman, R. B. (1999): "The economics of crime," *Handbook of labor economics*, 3, 3529–3571.

Gendreau, P., T. Little, and C. Goggin (1996): "A meta-analysis of the predictors of adult offender recidivism: What works!" *Criminology*, 34, 575–608.

Greene, W. H. (2003): *Econometric analysis*, Pearson Education India.

Hanson, R. K. (2002): "Recidivism and age: Follow-up data from 4,673 sexual offenders," *Journal of interpersonal violence*, 17, 1046–1062.

Institut de la statistique du Québec (2018a): "Indice d'ensemble des prix à la consommation," [Online; accessed 13-May-2018].

——— (2018b): "Population et structure par âge et sexe," [Online; accessed 7-May-2018].

——— (2018c): "Revenu disponible par habitant, régions administratives et ensemble du Québec, 2006-2017," [Online; accessed 14-May-2018].

——— (2018d): "Taux de chômage, régions administratives, régions métropolitaines de recensement et ensemble du Québec, 2007-2017," [Online; accessed 7-May-2018].

Jung, H., S. Spjeldnes, and H. Yamatani (2010): "Recidivism and Survival Time: Racial Disparity among Jail Ex-Inmates," *Social Work Research*, 34, 181–189.

Kennedy, P. (2003): *A guide to econometrics*, MIT press.

Kurlychek, M. C., R. Brame, and S. D. Bushway (2006): "Scarlet letters and recidivism: Does an old criminal record predict future offending?" *Criminology & Public Policy*, 5, 483–504.

——— (2007): "Enduring risk? Old criminal records and predictions of future criminal involvement," *Crime & Delinquency*, 53, 64–83.

Lin, M.-J. (2008): "Does unemployment increase crime? Evidence from US data 1974–2000," *Journal of Human Resources*, 43, 413–436.

McGovern, V., S. Demuth, and J. E. Jacoby (2009): "Racial and ethnic recidivism risks: A comparison of postincarceration rearrest, reconviction, and reincarceration among White, Black, and Hispanic releasees," *The Prison Journal*, 89, 309–327.

Messner, S. F. and R. J. Sampson (1991): "The sex ratio, family disruption, and rates of violent crime: The paradox of demographic structure," *Social Forces*, 69, 693–713.

Mitchell, O., J. C. Cochran, D. P. Mears, and W. D. Bales (2017): "Examining Prison Effects on Recidivism: A Regression Discontinuity Approach," *Justice Quarterly*, 34, 571–596.

Olver, M. E. and S. C. Wong (2015): "Short-and long-term recidivism prediction of the PCL-R and the effects of age: A 24-year follow-up." *Personality Disorders: Theory, Research, and Treatment*, 6, 97.

Piquero, A. R., W. G. Jennings, B. Diamond, and J. M. Reingle (2015): "A systematic review of age, sex, ethnicity, and race as predictors of violent recidivism," *International journal of offender therapy and comparative criminology*, 59, 5–26.

Report of the Auditor General of Quebec to the National Assembly for 2016-2017 (2016-2017): "Correctional Services: Reintegration Component," [Online; accessed 11-May-2018].

Rocque, M., C. Posick, S. E. Barkan, and R. Paternoster (2015): "Marriage and county-level crime rates: A research note," *Journal of Research in Crime and Delinquency*, 52, 130–145.

Schnepel, K. T. (2013): "Labor Market Opportunities and Crime: Evidence from Parolees," Tech. rep., Mimeo, University of California at Santa Barbara.

Siwach, G. (2018): "Unemployment shocks for individuals on the margin: Exploring recidivism effects," *Labour Economics*.

South, S. J. and K. M. Lloyd (1992): "Marriage opportunities and family formation: Further implications of imbalanced sex ratios," *Journal of Marriage and the Family*, 440–451.

Stephens, S., M. C. Seto, A. M. Goodwill, and J. M. Cantor (2016): "The relationships between victim age, gender, and relationship polymorphism and sexual recidivism," *Sexual Abuse*, 1079063216630983.

The Constitution Act (1867): [Online; accessed 11-May-2018].

Thornton, D. (2006): "Age and sexual recidivism: A variable connection," *Sexual abuse: A journal of Research and Treatment*, 18, 123.

Tripodi, S. J., J. S. Kim, and K. Bender (2010): "Is employment associated with reduced recidivism? The complex relationship between employment and crime," *International Journal of Offender Therapy and Comparative Criminology*, 54, 706–720.

Visher, C. A., L. Winterfield, and M. B. Coggeshall (2005): "Ex-offender employment programs and recidivism: A meta-analysis," *Journal of Experimental Criminology*, 1, 295–316.

Wooldridge, J. M. (2010): *Econometric analysis of cross section and panel data*, MIT press.