On Locus of Control in Empirical Microeconomics

Inaugural-Dissertation

zur Erlangung des akademischen Grades

einer Doktorin der Wirtschaftswissenschaft

vorgelegt am

Fachbereich Wirtschaftswissenschaft

der

Freien Universität Berlin

vorgelegt von

Juliane Hennecke (M.Sc.)

geborene Pohl

Berlin 2019

Erstgutachter:

Prof. Dr. Ronnie Schöb
(Freie Universität Berlin)

Zweitgutachter:

Prof. Dr. Marco Caliendo
(Universität Potsdam)

Tag der Disputation:

1. Juli 2019

Für Ulf, Lilli & Louisa

Erklärung zu Koautoren

Die vorliegende Dissertation umfasst eine Einleitung (Kapitel 1) und vier Forschungspapiere (Kapitel 2 bis 5). Die Kapitel 1 und 2 sind allein verfasst wurden. Die Kapitel 3 bis 5 sind in Ko-Autorenschaft entstanden. Ko-Autoren von Kapitel 3 sind Prof. Dr. Marco Caliendo, Prof. Deborah Cobb-Clark und Prof. Dr. Arne Uhlendorff. Prof. Dr. Marco Caliendo ist Ko-Autor von Kapitel 4 und Kapitel 5 ist in Zusammenarbeit mit Malte Preuß entstanden.

Für die Dissertation sind diese Kapitel gegenüber den gemeinsam verfassten Version redaktionell leicht angepasst worden. Diese Veränderungen verantwortet allein der Autor der vorliegenden Dissertation.

Vorveröffentlichungen

Die folgende Liste enthält alle Vorveröffentlichungen. Darunter sind auch Versionen der Kapitel, die zum Teil stark überarbeitet wurden, bevor sie Eingang in die vorliegende Dissertation fanden.

Kapitel 2: The Independent Woman - Locus of Control and Female Labor Force Participation

- *keine Vorveröffentlichungen*

Kapitel 3: Locus of Control and Labor Market Migration

- IZA Discussion Paper Series, No. 9600
- SOEPpapers on Multidisciplinary Panel Data Research 818, DIW Berlin

Kapitel 4: Drinking is Different! Examining the Role of Locus of Control for Alcohol Consumption

- *keine Vorveröffentlichungen*

Kapitel 5: Biased by Success and Failure: How Unemployment Shapes Locus of Control

- Labour Economics, Volume 53, pp. 63-74.
- School of Business and Economics Discussion Paper, Freie Universität Berlin, 2017/29
- SOEPpapers on Multidisciplinary Panel Data Research 818, DIW Berlin

Rechtliche Erklärung

Erklärung gem. §4 Abs. 2 (Promotionsordnung)

Hiermit erkläre ich, dass ich mich noch keinem Promotionsverfahren unterzogen oder um Zulassung zu einem solchen beworben habe, und die Dissertation in der gleichen oder einer anderen Fassung bzw. Überarbeitung einer anderen Fakultät, einem Prüfungsausschuss oder einem Fachvertreter an einer anderen Hochschule nicht bereits zur Überprüfung vorgelegen hat.

Berlin, März 2019
Juliane Hennecke

Erklärung gem. §10 Abs. 3 (Promotionsordnung)

Hiermit erkläre ich, dass ich für die Dissertation folgende Hilfsmittel und Hilfen verwendet habe: LaTeX, Stata, Mendeley, Literatur siehe Literaturverzeichnis. Auf dieser Grundlage habe ich die Arbeit selbstständig verfasst.

Berlin, März 2019
Juliane Hennecke

Acknowledgements

I would like to thank a number of people for their help and support, which was indispensable for the preparation of this dissertation.

Above all, I express my greatest thank to Ronnie Schöb, my primary supervisor. Whenever I had problems, questions, concerns or the craziest ideas, his door was open for me. I learned a lot about being a good economist from him as he always remembered me that doing research is all about making the world a little better by asking nasty questions about the real world implications and consequences of my work. He ensured all the freedom I required to develop my very own research agenda, while providing clear guidance whenever I needed it. Secondly, I would particularly like to thank Marco Caliendo. He recognized my passion and talent for economic research before I recognized it myself and kept on supporting me during all those years. I am thankful to him for sharing all his knowledge and expertise with me during our joint projects and for additionally also closely accompanying my other projects. Above all, I am more than grateful that he was willing to spontaneously becoming a member of my PhD committee just before the end. I also thank Jule Specht for being my second supervisor during all those years. She was my connection to the world of psychological research and always enthusiastic about my ideas and very open to the "girl from economics" doing research on psychology while not having a clue of the field.

During my PhD, I interacted with and benefited from many more inspiring and talented researchers. I thank Deborah Cobb-Clark and Arne Uhlendorff for their cooperation during my first careful steps into economic research. Much of the skills I acquired during my PhD resulted from our cooperation. Additionally, I am greatly indebted to Malte Preuß, a great colleague and even better co-author. Although we had a hard time coordinating our schedules and views on the world of economics, our paper ended up being a great piece of cooperative work not least because of his persistency and accuracy. In addition, I am grateful to many colleagues at Freie Universität Berlin and University of Potsdam for their various support and for accompanying me during the process. I thank, in particular, Clemens Hetschko, Daniel Nachtigall, Tobias Wolff, Cordula Arlinghaus, Martin Weißenberger, Linda Wittbrodt, Cosima Obst, Stefan Tübbicke, Mia Teschner and Claudia Stier. Furthermore, I am grateful for the financial support of the German Science Foundation (DFG) through the

Research Training Group (RTG) 1659 "Interdependencies in the regulation of markets" at the Humboldt-Universität zu Berlin as well as all my fellow students in the doctoral program for helping me survive the first year of courses and always standing by my side with ideas, help and support during the following years. I additionally thank Julia Hennecke, Luise Stürmer and Kira Güther for their support during the proof-reading phase in the past weeks.

While writing this thesis, I experienced ups and downs. I was however always sure to find support and encouragement in my family and friends for which I am grateful. One of the most important thanks goes to my parents Wolfgang und Kerstin. No words can express how important you are for my life and how happy I am to make you proud. Because of your love and support I became who I am today. Finally, the biggest and most important thank goes to my wonderful husband Ulf and our beautiful daughters Lilli and Louisa. Ulf, without your unrestrained and infinite support, not one of the words on the following pages would have been possible. Thank you for bearing with my bad moods if the regressions did not work, cooking for me when I was to deepened into my work, staying at home with Lilli and Louisa when I traveled the world or spend long evenings in the office and for drying my tears when the despair overwhelmed me. I love you. Lilli and Louisa, your smiles are the one and only thing on this world that keep me going every day. You are our greatest gifts and the loves of my life.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Locus of Control in Empirical Microeconomics - An Introduction

"The foundation of political economy and, in general, of every social science, is evidently psychology. A day may come when we shall be able to deduce the laws of social science from principles of psychology [...]."
(Pareto, 1906, p.20)

1.1 Personality Psychology and Behavioral Economics

Interdisciplinarity can, without a doubt, be named as one of the major trends in modern economic research. Although it was common practice in other disciplines since their day one and also economics was rich of interdisciplinary ideas in its early years, for a long time the majority of economists was certain that they could explain the world's states by solely relying on their own models and theories. Understanding and predicting human behavior, thus, was widely restricted to fitting the behavior of humans into the concept of a perfectly informed, fully rational, utility maximizing homo economicus. Although this "Econ" is without a doubt the greatest achievement of microeconomic theory, economists were criticized for their simplification of the world mainly because the lines between normative and positive economics started to blur. In the words of Richard Thaler, economists forgot that *"... we do not live in a world of Econs. We live in a world of Humans"* (Thaler, 2015, p.16). For many decades, economists ignored that the original idea of the homo economicus was to describe how humans should logically behave instead of how they actually behaved. Thus, for a long time individual deviations from the rational choice and utility maximization frameworks were attributed to idiosyncratic shocks and added into error terms. (Thaler, 2015)

Nevertheless, especially the rise of modern behavioral economics has tremendously changed the economic discipline. Economists started to differentiate between normative and positive approaches. Especially empirical microeconomics is now largely motivated by describing human behavior and decision making as accurately as possible and thus identifying whether and why the homo sapiens does deviate from the theoretical prescriptions for the homo economicus. Thanks to this return to the traditional virtues of the economic theory, the concept of economic actors being "Econs" can unfold its full potential of being a powerful guideline for the analysis of individual behavior throughout the life-cycle. Empirical economic research which is guided, but not driven, by these underlying core concepts of economics while at the same time taking into consideration most recent evidence from all other social as well as natural sciences has a tremendous potential of getting one step closer to understanding human behavior.

Especially psychological theories and ideas have enriched economic research in the past decades. The Nobel laureates Daniel Kahneman and Richard Thaler as well as their various colleagues made psychologically motivated theory fit for economic models. Nowadays, concepts and models from psychology are ubiquitous in economics. One of the most prevalent psychological concepts in modern empirical microeconomics are personality traits and their influence on economic decision making and economic outcomes. In line with the well-known definition by Roberts (2009, p.140), personality traits can be defined as *"...relatively enduring patterns of thoughts, feelings, and behaviors that reflect the tendency to respond in certain ways under certain circumstances"*. Although this definition is widely spread and largely accepted, the literature is more than discordant on what those patterns actually include. In

line with Borghans *et al.* (2008, p.974) most economists nevertheless assume that personality traits concern *"all dimensions on which people differ from one another"*.[1] In addition to cognitive skills, which mainly describe an individuals problem-solving abilities, economists often assume personality traits to be a bundle of all inherent skills which are not directly related to knowledge and ability and thus often label them as "non-cognitive skills". (Borghans *et al.*, 2008) Almlund *et al.* (2011, p.5) describe personality as a *"strategy function for responding to life situations"*.

In the past few years, personality traits have increased in importance for the explanation of various types of individual behavior and decision making in microeconomic studies. The increasing availability of survey-based measures of psychological concepts made it more and more attractive for applied economists to investigate this psychological black box behind human behavior. Most notably thanks to the work of James Heckman and his various coauthors, modern economists do not have doubts about the important predictive power of these non-cognitive skills for economic outcomes which has been found to exceed that of cognitive abilities (see e.g. Heckman *et al.*, 2006; Heckman and Rubinstein, 2001). Especially, their extensive research on the famous Perry Preschool Program[2] was a milestone in demonstrating that these traits actually do greatly matter for individual economic outcomes (Heckman *et al.*, 2013). Motivated by their seminal work, the economic literature is continuously growing and has found important effects of personality traits e.g. on earnings as one of the hot topics within economics (Mueller and Plug, 2006; Nyhus and Pons, 2005).

Based on the fundamental ideas from traditional economic theory as well its extensions through modern behavioral economics, individual differences in decision making are largely simplified to originate in three important idiosyncratic components, which are 1) constraints and opportunities, 2) preferences as well as 3) beliefs and expectations. (Gintis, 2006) In line with the growing motivation to understand the psychological black box behind individual decision making, personality traits have been empirically found to be important determinants of all three aspects and thus are important components of choice models in all sorts of decision making situations (Borghans *et al.*, 2008; Almlund *et al.*, 2011).

Exactly this is the point at which this thesis makes its contribution. It presents three studies (Chapters 2-4) which consider one specific taxonomy of personality, the so called locus of control and its important role for decision making in three distinct domains of the life-cycle which are labor market mobility, labor force participation as well as risky health behavior. In doing so, it contributes to the growing empirical literature on the tremendous importance

[1] For more details on the psychological literature on aspects and dimensions of personality see e.g. McAdams and Pals (2006) and Roberts *et al.* (2006a).

[2] The Perry Preschool Program was a large randomized control trial targeted at disadvantaged African American pre-school children with low cogntive skills in the US in the early 1960s. The children were followed through the age of 40 which enabled researchers to identify effect of the intervention on earnings, education, crime and many other socio-economic outcomes. Over the years dozens of studies have analyzed the returns of the program. See e.g. Heckman *et al.* (2010) for an overview over the literature.

of this trait for individual behavior and economic decision making. Additionally, Chapter 5 crucially adds to the literature, by discussing the most prevalent problem in the literature on the predictive power of personality traits for individual decision making: the danger of reverse causality and attenuation bias through issues with the stability and measurement validity of personality traits.

The further parts of this introduction are structured as follows. Section 1.2 gives a detailed introduction into the theory behind locus of control and the psychological literature on the concept as well as the survey-based measurement of the trait. Subsequently, Section 1.3 introduces the previous literature on behavioral implications of locus of control. Finally, Section 1.4 summarizes the contributions of the thesis as well as the single studies.

1.2 A Primer into Locus of Control

The facets and aspects of individual personality are highly complex. The most prominent taxonomy of traits is the so-called Big Five which consist of the traits openness, conscientiousness, extraversion, agreeableness and neuroticism (McCrae and Costa Jr, 2008; John and Srivastava, 1999). Those traits are widely assumed to cover the majority of inter-individual differences in inherent characteristics. Nevertheless, during the past years, another important trait has made its way into the center of attention especially in empirical economics: the theory of internal versus external control of reinforcement, or locus of control (LOC).

1.2.1 The Taxonomy of Internal versus External Control of Reinforcement

The most prevalent definition of locus of control is still the one given in the seminal work by Julian B. Rotter. According to this, locus of control is the *"generalized attitude, belief, or expectancy regarding the nature of the causal relationship between one's own behavior and its consequences"* (Rotter, 1966, p.2). The basic idea emerged from Rotter's social learning theory (Rotter, 1954) which assumes a fundamental relationship between previously experienced reinforcements, expectations about future reinforcements and current behavior. According to Rotters theory, individuals experience rewards and punishments to their behavior. These reinforcements then again strengthen their expectations on future consequences to their own actions. Nevertheless, it is likely that *"...an event regarded by some persons as a reward or reinforcement may be differently perceived and reacted to by others."* (Rotter, 1966, p.1). This individual difference in the subjective expectations about *"the nature and effects of reinforcement"* (p.2) is resumed by an individual's perception of control. Rotter assumed that these expectancies are not only specific to certain situations but can very likely be generalized over large sets of individual behavior and are thus a dispositional trait. The locus of control, as it is largely called in modern psychological and economic research, thus describes an individual's subjective belief about the causal relationship between own efforts and abilities and life's outcomes. Consequently, it describes the location ("locus") of control being either

internal or external. Individuals who believe that future outcomes are mainly driven by own efforts and abilities have an internal locus of control as they perceive their individual control to be rather high. By contrast, individuals who mainly attribute life's outcomes to external factors such as fate, luck, a higher power or other people have an external locus of control as they perceive their individual control to be low. Although Rotter's article was not the first which attempted to formalize and measure perceptions of control (see e.g. Phares, 1957; James, 1957), it was especially in the years after Rotter's seminal work that locus of control has become an inherent part of the core concepts of personality psychology and the work of Rotter (1966) as well as a the prominent review article by Lefcourt (1966) became two of the most cited articles in the field (Lefcourt, 1992). Over the years, the general idea of control beliefs has been extended to various concepts of domain-specific perception of control such as e.g. in health (Wallston *et al.*, 1976), politics (Davis, 1983) as well as economics and finances (Furnham, 1986).

1.2.2 Interdependencies with Other Constructs

Although there was also considerable criticism especially concerning the measurement of the construct (see e.g. Brewin and Shapiro, 1984), locus of control gained more and more popularity with every study which was able to identify important psychological and behavioral consequences such as e.g. for motivation or self-control (see e.g. Skinner, 1996; Lefcourt, 1982; Rosenbaum, 1980). Nevertheless, an important discussion which has not come to an end yet concerns the unique feature of locus of control and the independence of the construct from other prominent taxonomies of traits. For example, Skinner (1996) was able to identify more than 100 constructs which reflect aspects of perceived control such as e.g. mastery, helpless and autonomy. Throughout the literature, locus of control is quite frequently related to prominent concepts such as self-esteem, self-efficacy and confidence (Hill, 2011). Also, the question of whether LOC is only a sub-trait of the Big Five as the taxonomy which captures all core facets of individual personality is still heavily discussed. (Bouchard and Loehlin, 2001). Especially Timothy A. Judge and his colleagues discussed the interdependencies between locus of control and other constructs such as self-esteem, self-efficacy and emotional stability (i.e. neuroticism) by proposing an underlying lower-dimensional factor called "core self-evaluation". (Judge *et al.*, 2002; Judge and Hurst, 2007; Bono and Judge, 2003; Judge *et al.*, 2003) Similar attempts to relate locus of control especially to the concept of self-efficacy have been made e.g. by Pearlin and Schooler (1978) and Ajzen (2002). Although they are closely related, a distinct difference between the concepts of self-efficacy, self-esteem and locus of control can be stated. As summarized by Cobb-Clark (2015), locus of control refers to the causal relationship between own behavior and life's outcomes in general whereas self-efficacy rather describes the individual belief in the personal ability to be capable of acting effectively to bring about desired results in a specific task. Thus, while self-efficacy describes whether an individual thinks that she is able to perform a certain task, locus of control describes whether successfully

performing a task will affect future outcomes after all. Self-esteem again rather describes the overall confidence in own abilities in general. According to Hill (2011), the crucial difference to locus of control lies in the processing of negative reinforcements. Thus, individuals with an internal locus of control might very well develop a low self-esteem as they also directly attribute negative reinforcements to their own actions and thus failures.

1.2.3 Measurement and Scales

Psychological Tests and Survey Measures In line with other personality concepts, locus of control, too, is traditionally measured using a variety of tests and self-reports with the basic idea of identifying individual differences. (Borghans *et al.*, 2008) Rotter and his colleagues elaborated a 29-item forced-choice test in order to obtain a quantitative measure of the underlying construct. In their list of 29 pairs of statements, they thus asked survey respondents to e.g. decide whether "Becoming a success is a matter of hard work, luck has little or nothing to do with it" or "Getting a good job depends mainly on being in the right place at the right time." (see Rotter, 1966, p.11) Since then, a vast number of different scales have been developed in order to capture this latent trait within surveys and psychological tests (see e.g. Lefcourt, 1982, for an comprehensive overview). Nevertheless, while the main goal of these scales developed by psychologists is to draw a picture of the trait which is as accurate as possible, in order to isolate it from other related facets of the individual personality, empirical economists are rather interested in the practicability of the measurement in the context of large-scale surveys. As surveys are limited in the array measures they include, the survey-based measurement of personality traits in general and locus of control in specific is extremely restricted to a small, reduced set of questions or items (Borghans *et al.*, 2008). As is comprehensively described by Cobb-Clark (2015), there is thus a distinct tradeoff between getting to the core of the concept one wants to analyze and nevertheless being able to utilize survey-based measures, which is nevertheless largely ignored in *"economist's laissez-faire approach"* (p.3). Yet, a growing psychological literature on the reliability of reduced-set survey-based measures (see e.g. Roberts *et al.*, 2011) additionally increased the attractively of these concepts to economists in the past years. In most surveys, respondents are asked to rate their agreement with a series of statements on Likert-scales ranging from strong disagreement to strong agreement.

International Surveys and their Current Locus of Control Scales Nevertheless, the modules which are assumed to capture locus of control still noticeably differ between surveys. Herein lies an important potential source of differences in the empirical results based on the different surveys. Table 1.1 gives an overview over the available modules in three of the largest international surveys.

Researchers working with the Australian "Household, Income and Labour Dynamics in Australia (HILDA)" solely make use of items from the Mastery Module developed by Pearlin

Table 1.1: Locus of Control Scales in International Surveys

SOEP	HILDA	NLSY79
(1) How my life goes depends on me	(1) I have little control over the things that happen to me	(1a) What happens to me is my own doing
(2) Compared to other people, I have not achieved what I deserve	(2) There is really no way I can solve some of the problems I have	(1b) Sometimes I feel that I don't habe enough control over the directions my life is taking
(3) What a person achieves in life is above all a question of fate or luck	(3) There is little I can do to change many of the important things in my life	(2a) When I make plans, I am almost certain that I can make them work
(4) If a person is socially active, she can have an effect on social conditions	(4) I often feel helpless in dealing with the problems of life	(2b) It is not always wise to plan too far ahead, because many things turn out to be matter of good or bad fortune anyhow
(5) I have the experience that others have a controlling influence over my life	(5) Sometimes I feel that I am being pushed around in life	(3a) In my case, getting what I want has little or nothing to do with luck
(6) One has to work hard in order to succeed	(6) What happens to me in the future mostly depends on me	(3b) Many times we might just as well decide what to do by flipping a coin
(7) If I run up against difficulties in life, I often doubt my own abilities	(7) I can do just about anything I really set my mind to do	(4a) Many times I feel that I have little influence over the things that happen to me
(8) The opportunities that I have in life are determined by the social conditions		(4b) It is impossible for me to belief that chance or luck plays an important role in my life
(9) Inborn abilities are more important than any efforts one can make		
(10) I have little control over the things that happen in my life		

Source: SOEP waves 1999,2005, 2010 and 2015, HILDA 2003, 2004, 2007 and 2011 and NLSY79 1976 and 2014.

and Schooler (1978) in order to capture locus of control. As this module by definition has the intention to capture self-efficacy rather than locus of control, these studies largely abstract from the differences and rather rely on the similarities between the concepts. (see e.g. Cobb-Clark *et al.*, 2016; Cobb-Clark and Schurer, 2011; Cobb-Clark *et al.*, 2014) As opposed to this, the "National Longitudinal Survey of Youth 1979 (NLSY79)" also captures the Pearlin Mastery scale as a measure for self-efficacy but additionally comprises a four-item abbreviated version of the original forced choice questionnaire by Rotter as a measure of locus of control (see NLSY79, 2019, for an overview). The measurement of locus of control in the German Socio-Economic Panel (SOEP), which will also be the basis of the studies presented in the following chapters, is based on a psychometric scale specifically developed for the SOEP by Helmut Nolte based on the evaluation of an older set of items in Nolte *et al.* (1997). The developed scale was, in the original conception, intended to cover internal and external locus of control but additionally also attitudes towards fairness or justice (item 2) as well as individual vs. collective orientation (item 4) with one item each (Richter *et al.*, 2013). The list contains statements which are very close to the original statements by Rotter but some items are also very similar to those contained in Pearlin's Mastery module. The literature using this measure is relatively disjoint about the question which items should be used for the construction of the locus of control scale. In line with Specht *et al.* (2013), at least seven out of the ten items can be combined into an overall scale with good reliability. While Specht *et al.* (2013) and Richter *et al.* (2013) suggest to exclude items 4, 6 and 9, many studies rests upon the original concept by excluding item 2 and 4 from the list (see e.g. Pinger *et al.*, 2018). In addition to this multiple other combinations of included and excluded items exist in the literature (see e.g. Berger and Haywood, 2016; Schnitzlein and Stephani, 2016; Caliendo *et al.*, 2014; Heywood *et al.*, 2017).

Factor Analysis and Latent Traits Most of these studies, including those presented in the following chapters, base their decision about included and excluded items upon exploratory factor analysis. Independent from the SOEP and the choice of items, also Borghans *et al.* (2008, p.978) note that the concept of factors is *"...(c)entral to [...] recent empirical work at the intersection of economics and psychology"*. Factor analysis has the goal of identifying clusters within the surveyed modules in order to isolate quantitative measures of the latent factors underlying the theoretical constructs. See for example Borghans *et al.* (2008) for a comprehensive introduction into factor analysis as well as a discussion of the importance of the concept of factor for the interface between personality psychology and economics. While the earlier literature on economic consequences of locus of control largely abstracted from the importance of factor analysis and the resulting loadings for the construction of the latent trait, especially the more recent studies now rely on confirmatory factor analysis in order to predict latent traits. This was mainly motivated by the work of Piatek and Pinger (2016), who discuss important issues with attenuation bias if a simple average with equal weights is used

instead of the predicted factor, which allows for a sample-driven weighting of the included items.

1.3 Previous Literature - Locus of Control, Behavior and Economic Outcomes

The increasing availability of survey-based measures of locus of control especially in large longitudinal household studies in Germany and Australia has led to a large body of studies which attempt to identify the predictive power of locus of control for individual decision making and behavior in various domains.

In this literature, locus of control is widely assumed to be an incentive-enhancing preference, in line with the theoretical discussion by Bowles *et al.* (2001). As an increased internal locus of control is associated with an increased belief in the impact of own actions on own outcomes, most of the existing literature theoretically connect locus of control and individual behavior through the enhancement of efforts and investments into future outcomes.

It was especially in the past two decades, that economists have begun to explicitly incorporate locus of control as well as other personality traits into their economic models of decision making in order to understand why they seem to matter for the behavioral outcomes we observe (see e.g. Cobb-Clark, 2015; Almlund *et al.*, 2011; Borghans *et al.*, 2008). Much of the growing strand of literature on the behavioral implications of locus of control was initiated and motivated by the work of Coleman and DeLeire (2003) on the effect of LOC on human capital investments. Not only were they the first who analyzed how LOC affects educational outcomes, but their study was a seminal work for the theorisation of the effect of LOC on beliefs about returns to investments in line with the argument about incentive-enhancing preferences proposed by Bowles *et al.* (2001). In their model, they argue that individuals with an internal locus of control have stronger beliefs into the effect of their educational investments on future wages and thus invest more. They, additionally, provide empirical evidence for their arguments using data on US adolescence. Also Heckman and Kautz (2012) and Mendolia and Walker (2014b) find important effects of locus of control on educational attainment in adolescence and Caliendo *et al.* (2016) and well as Offerhaus (2013) add evidence for educational investments in adulthood.

Large parts of the earlier economic literature, which related economic outcomes, such as wages, and locus of control attributed the effects to differences in levels of non-cognitive skills (Andrisani, 1977; Duncan and Dunifon, 1998; Goldsmith *et al.*, 1997). As opposed to this, the literature which emerged after the work of Coleman and DeLeire (2003) also consider behavioral implications of LOC such as e.g. differences in the investment into human capital, as indirect drivers of observed differences in the economic outcomes (Heineck and Anger, 2010;

Piatek and Pinger, 2010; Semykina and Linz, 2007; Groves, 2005; Schnitzlein and Stephani, 2016).

Not only did the work of Coleman and DeLeire (2003) change the lines of arguments in the literature on wage effects of locus of control, but it also launched a new strand of literature, which focused on the effect of LOC on investments into future economic outcomes. Caliendo *et al.* (2015b), McGee (2015) and McGee and McGee (2016), for example, discuss and analyze the effect of LOC on job-search behavior via different subjective beliefs about returns to search. In line with these findings, studies by, for example, Uhlendorff (2004) as well as Gallo *et al.* (2003) provide evidence that internal job seekers indeed have higher reemployment probabilities and shorter unemployment periods on average.

Moverover, multiple studies also relate locus of control and occupational choice and attainment. Heywood *et al.* (2017) provide evidence that internal individuals sort into jobs with performance appraisals as they appreciate that these jobs allow them to demonstrate their value and achieve their goals. Closely related to the existing evidence by Andrisani (1977), Cobb-Clark and Tan (2011) find that especially internal men are on average more likely to sort into jobs with personnel responsibility, i.e. they are significantly more likely to be in managerial positions, and also more likely to be education professionals. Eventually, Ahn (2015) adds empirical evidence that individuals with an internal LOC are on average observed to perform job-to-job transitions which improve their wages more often. Furthermore, internals have not only been shown to be more likely to become self-employed but they are additionally also more successful entrepreneurs. (Caliendo *et al.*, 2014; Hansemark, 2003; Rauch and Frese, 2007). In accordance with all these results, Ng *et al.* (2006) conduct a meta-analysis based on which they are able to conclude that an internal LOC improves individuals job satisfaction and job motivation.

However, locus of control has not only been related to individual behavior and outcomes on the labor market. For example Cobb-Clark *et al.* (2014), Mendolia and Walker (2014a) and Chiteji (2010) provide evidence that locus of control and the related concept self-efficacy are associated with less risky and more conscious health behavior. Additionally, Lekfuangfu *et al.* (2018) find an important effect of LOC on parental investments. Also the role of locus of control for behavior on the financial market has been repeatedly analyzed especially in the past years. Hence, Cobb-Clark *et al.* (2016) find that households with an internal reference person on average have higher savings rates while Salamanca *et al.* (2016), Pinger *et al.* (2018) and Kesavayuth *et al.* (2018) relate an internal locus of control and financial investment decisions and financial risks. Salamanca *et al.* (2016) show that internal individuals are more likely to invest into risky assets and Kesavayuth *et al.* (2018) support their findings by showing that internal individuals are in general more willing to take financial risks. As opposed to the pervasive mechanism of differences in expectations about future returns, both studies argue that a likely mechanism behind an effect of LOC on risky financial behavior are differences

in the risk perceptions. Individuals who have an internal LOC simply perceive less variance in risky assets because they believe that, through their own actions, they are more in control of future outcomes compared to those with an external LOC. In line with this, Pinger *et al.* (2018) provide experimental evidence that internal individuals are also more likely to make inconsistent risk choices and more likely to fall for the hot hand fallacy. As is notes by Cobb-Clark (2015), this new strand of literature adds an interesting new perspective to the role of LOC for economic decision making as it proposes an alternative mechanism which can also be applied to other investment decisions such as human capital and health as also these investment decisions are associated to risk and uncertainty.

As a third highly relevant mechanism for economic behavior and outcomes, Buddelmeyer and Powdthavee (2016) as well as Schurer (2017) find an important insurance effect of LOC against negative life events such as in specifically health shocks. While Buddelmeyer and Powdthavee (2016) show, that internal individuals experience a less severe reduction in life satisfaction as a reaction to e.g. own and and others' serious health shocks as well as crime, Schurer (2017) provides evidence that individuals with an internal LOC are less likely to drop out of the labor force after they experienced a health shock.

A comprehensive overview over the literature on locus of control in labor economics is also given in Cobb-Clark (2015). A more detailed discussion of the existing literature in the specific domains of female labor force participation, migration decisions as well as risky health behavior are presented in the literature reviews of Chapters 2, 3 and 4. The growing body of literature on the stability of personality traits is discussed in detail in the literature review of Chapter 5.

1.4 This Thesis - Contributions and Main Findings

1.4.1 Behavioral Implications of Locus of Control (Chapter 2-4)

In line with the increasing awareness for the role of personality traits for economic decision making, the present thesis makes its core contribution especially to the growing empirical literature on the explanatory power of the personality trait locus of control for individual behavior. It does so by presenting three studies (Chapters 2-4) on the important role of LOC for decision making in three distinct domains of the life-cycle which are labor market mobility, labor force participation as well as risky health behavior.

Besides the focus on locus of control, the distinct feature of all three studies is that they not only identify the empirical relationship between locus of control and the concerned decision situation but they also make a considerable effort to carefully introduce locus of control into the prevalent economic theory behind each of these choice situations. Hence, it is the common goal of all chapters in this thesis to consider the psychological roots of the concept, understand the meaning behind it and try to identify the channels of its virtue on

economic decision making. All studies make an attempt to pin down the knowledge gains and implications of the findings for traditional economic decision making models by carefully discussion channels and mechanism within the framework of rational choice theory. The most prevalent common finding of all studies is that the relationship between personality traits and economic behavior is a highly complex network of simultaneous impacts which is additionally highly heterogeneous between situations and individuals. All chapters thus share not only a common explanatory variable of interest as well as a common data basis, which is the German Socio-Economic Panel (SOEP), but also have a very similar scope as well as theoretical and empirical approach.

In line with the fifth chapter, all parts of the thesis additionally share their careful consideration of potential risks from reverse causality through doubts about the stability of locus of control as well as its measurement validity. This issue is "the elephant in the room", present in all chapters of the thesis. Besides a careful discussion of potential biases in the empirical sections of Chapters 2 to 4, Chapter 5 completes the picture by carefully discussion the danger of reverse causality and attenuation bias through issues with the stability of personality traits and measurement validity in the specific case of exogenous unemployment shocks. Additionally, the implications of Chapter 5 for the other chapters are carefully discussed within the chapters as well as in section 1.4.2 of this introduction. Consequently, the empirical approaches of this thesis largely profit from the synergy between economic and psychological models. The integration of the empirical findings into the more fundamental principles and theories of economics and psychology is a witness of the important interrelationship between both disciplines.

Female Labor Force Participation (Chapter 2)

Chapter 2 initiates the discussion by analyzing the implications of LOC for female labor force participation. Research on female labor force participation has a long tradition in theoretical and empirical economic research. While a lot of open questions have been answered on the gender gap in labor market participation, the prevalent heterogeneity between women still keeps economists and politicians busy. While traditional economic theory attributed unexplained differences in decision making to idiosyncratic shocks, more modern approaches are more and more informed by psychology and thus consider individual differences in tastes, preferences and beliefs as important drivers of decision making. In line with these advancements in the literature, the paper theoretically discusses the role of locus of control for participation decisions by incorporating a vector of individual attributes into a woman's utility and budget function which captures LOC. The vector of attributes is argued to affect the preferences for leisure, home production and childcare as well as participation but also the expectations about monetary returns to participation. In line with the psychological literature, especially

a crucial importance of LOC for independence preferences as well as subjective beliefs about returns to effort and investments are proposed to be important mechanisms.

The empirical analysis supports these theoretical considerations. Using data from the SOEP, internal women are found to have a significantly higher probability of being available for market production, which also translates into higher employment probabilities at the extensive margin. These effects are additionally found to be highly heterogenous with respect to underlying monetary incentives for participation and home production as well as prevalent social norms for working.

Hence, the paper not only contributes to the literature on the economic importance of LOC as well as the psychological black box behind participation decisions but is also able to more generally identify important interdependencies between monetary constraints and opportunities as well as inherent preferences and attributes in the decision making process.

Internal Migration (Chapter 3)

In a quite similar manner, Chapter 3 discusses the role of LOC for regional labor market mobility within Germany. The study theoretically models internal migration based on job seekers who form subjective beliefs about the return to search effort that are related to their locus of control. In line with the findings in Caliendo *et al.* (2015b), internal job seekers are predicted have higher subjective expectations about the returns to their search efforts and thus search more intensively. Given the relationship between search effort and geographic search area, this implies that individuals with a more internal locus of control will send more applications to other geographic regions than individuals with a more external locus control. These considerations are supported with an ancillary analysis using data from the IZA Evaluation data which identifies a significant positive relationship between having an internal LOC and applying for jobs for which individuals would have to move. In a situation in which job search precedes migration decisions, this increased geographic distant search is assumed to lead individuals to migrate more frequently. Nevertheless, if job search occurs post-migration, locus of control and internal migration are expected to be unrelated.

The implications of the model are tested using SOEP data via which the exact region of residence can be identified for every household. The empirical analysis identifies a distinct positive effect of an internal LOC on the general self-reported willingness to move as well as the probability of moving between regions.

Consequently, also this paper not only contributes to the literature on the behavioral implications of LOC and the psychological determinants of migration decisions but also makes a more fundamental contribution to the literature by demonstrating that assumptions about the timing of job search have important implications for the pattern of internal migration that results.

Alcohol Consumption (Chapter 4)

A prove that the importance of LOC for decision making cross the boarders of labor economics is provided within Chapter 4. The chapter is devoted to the question of whether LOC is also able to explain alcohol consumption as an important domain of risky health behavior. The existing empirical literature on alcohol consumption shows that drinking is linked to extreme risks for the consumer's physical and mental health and thus preventable micro- and macroeconomic costs such as e.g. through the strain it puts on individual health care expenditures and labor market perspectives. Nevertheless, it is at the same time associated with a high level of uncertainty about future consequences, as moderate levels of consumption have frequently been related to positive outcomes for health and the economic situation.

By making use of the unique information available on drinking behavior within the SOEP, the study identifies a strongly positive effect of an internal LOC on moderate as well as heavy drinking which is comparable to effect of traditional preference parameters such as risk aversion and time preferences. The paper suggests and discusses two likely mechanisms for this relationship. On the one hand, internal individuals are expected to put a higher effort into the investment in social and occupational networks and via this have a higher likelihood to drink as the attendance of social gatherings is often inextricably linked with alcohol consumption. On the other hand, as uncertainty and the perception of risks play a crucial role for individual considerations about alcohol consumption, internals consume more alcohol on average as they believe in or overestimate their personal ability to cope with or prevent the negative consequences of consumption. Based on a number of ancillary analyzes, the paper conclude that very likely both channels are at play.

Hence, also this paper again contributes to the literature on behavioral implications of LOC as well as to the explanation of heterogeneity in individual decision making in a domain with fundamental economic importance. As opposed to large parts of the existing literature on the behavioral consequences of LOC also in the area of health, the study is novel in a sense that it is able to identify potentially negative consequences of an internal LOC. Always striving for a LOC which is as internal as possible might thus not be the universal solution for social planners.

1.4.2 The Stability and Exogeneity of Locus of Control (Chapter 5)

Eventually, Chapter 5 importantly contributes to the value all parts of this thesis add to the body of literature on behavioral consequences of LOC by carefully discussing the stability of the trait and thus potential problems with reverse causality. As these are without a doubt "the elephant in the room" in the three other studies as well as in the existing literature on behavioral implications of personality traits, the methodological contribution is fundamental. The discussion on the stability of personality traits in general and locus of control in specific, is one of the central topics of personality psychology since multiple decades and still divide the

minds. More strikingly, although for example Specht *et al.* (2013) and Roberts *et al.* (2006b) found important variation of LOC and other personality traits with age also in adulthood, the early economic literature uniformly assumed exogeneity of LOC in their economic models. It were especially the studies by Cobb-Clark and Schurer (2011, 2013) which motivated economists to carefully reconsider the stability of those traits. In order add to the assessment of the stability of LOC, the study, on which Chapter 5 is based, investigates the reaction of reported LOC to an exogenous unemployment shock. Reassuringly, the empirical analysis finds no long-lasting effects of job losses due to plant closures on LOC and thus cannot reject the common assumption of its stability during adulthood. Hence, concerns with respect to endogeneity of LOC within the models and thus reverse causality can be alleviated. Nevertheless, the study identifies an important temporary deviation in the measurement of LOC during periods of unemployment and therefore concludes that the reported LOC is affected by unemployment likely due to a situation-specific effect. The effect holds true independent from unemployment duration or socio-demographic characteristics and vanishes as soon as the unemployed find a new job.

The study discusses these identified effects within the well-known trait-state model from personality psychology. The model accounts for two components: the time-invariant 'trait' and the situational characteristics the measurement takes place in, the so called 'state'. While the trait is – in line with the economic definition of personality – a highly enduring or even unchanging entity, the state is a temporary deviation from it, caused by the interaction between the survey participant and her current surroundings. Consequently, it is concluded that, rather than being concerned with reverse causality, the future literature needs to carefully consider potential measurement issues within LOC. Using the stated LOC as explanatory variable can lead to biased estimations when the temporary deviation in measurement is not accounted for. Nevertheless, it is beyond the scope of the study to identify wether the observed changes in the stated LOC do have behavioral implications or are pure measurement error.

Remark - Implications of Chapter 5 for the Findings in the Previous Chapters The implications of these conclusions for the findings in the other three chapters strongly depend on the nature of the identified state effects. When the state has the same properties as a measurement error, any estimations which pool individuals with different states will suffer from an attenuation or endogeneity bias depending on the correlation between its causing event and the outcome of interest (Bound *et al.*, 2001). If, on the other hand, the state effect has its own behavioral implications, the reported LOC correctly reflects behavior at any time. In this case, estimations will be biased as soon as the modeled outcome and the LOC measurement do not take place within the same state.

The appropriate econometric solutions for both cases therefore stand in sharp contrast. While the LOC which is observed at time t is the appropriated measure in the second case, using a lagged LOC which is observed in the same state for all individuals in the sample will

be sufficient to prevent any harm in the first case. Nevertheless, in addition to the discussion about employment states in Chapter 5, this might be true for many different states also related to other important domains such as family status or health. Observing all individuals in the sample during an exactly identical state thus is not feasible.

The chapter which is affected the most by those considerations is Chapter 2. Due to the fact that labor force participation as the outcome of interest is at risk of being the causing event for a potential measurement error in LOC, the findings are in danger of being biased upwards by endogeneity if women who do not participate in the labor force report lower levels of LOC and not vice versa. Consequently, this issue is carefully discussed within the study itself. In a robustness check, the lagged LOC from the last employment spell as well as an averaged LOC over all available observations are used as alternative measures in order to assess the severity of the bias (see Table 2.A.5). Reassuringly, the findings are robust against these adjustments. The risk that the conclusions of the study are affected by the findings in Chapter 5 is thus rather low.

Large parts of the study, on which Chapter 3 is based, have evolved and were written before the implications of Chapter 5 became apparent. This is why the implications of Chapter 5 are not discussed in detail within the chapter. Nevertheless, as opposed to Chapter 2, the risk of endogeneity bias is less strong due to the fact that LOC is always measured before migration decisions and employment states are included as control variables. A robustness check which accounts for earlier moves as a potential omitted variable which affects LOC as well as migration decisions reveals no problems in this respect too (see Table 3.4). As opposed to this, the findings might rather suffer from attenuation bias caused either by the fact that individuals are not necessarily observed during the same state or that LOC is not observed in exactly the same situation in which the decision making takes place depending on whether the measurement variability has behavioral implications. Reassuringly, a sensitivity check which uses the first LOC observation in the data (see Table 3.4) as well as an averaged measure of LOC over all observations demonstrate the robustness in this respect.

The issue is arguably the least severe for the study presented in Chapter 4. This is based on the assumption that different than labor market participation and regional mobility, alcohol consumption is no event which can be argued to have severe effects on the measurement of LOC if we assume that being drunk during the interview is a rather rare exception. Also in this study employment states are controlled for. Nevertheless, state effects in the measurement of LOC can still affect the estimated results through attenuation. Consequently, the problem is discussed within the sensitivity checks. An averaged measure as well as a measure which is imputed from the first available observation are used without any indication for severe biases (see Table 4.A.4).

The main takeaway from all these discussions and sensitivity checks nevertheless is that many, in parts counteracting, mechanisms are at play in the measurement of personality traits

leading to the conclusion that the perfect timing for the measurement of LOC probably does not exist. Depending on the nature of measurement errors, the discussed outcome variable and the considered sample, a prevalent conflict of interests exists. It is, on the one hand, important to measure LOC as close as possible to the decision situation in order to capture the trait which actually affected the individual at time t. On the other hand, lagging LOC might be an appropriate way of preventing endogeneity biases. Consequently, the most appropriate solution to this problem is a careful discussion of the issues and the use of a number of different measures in order to assess the extent of the problem.

CHAPTER 2

The Independent Woman - Locus of Control and Female Labor Force Participation

2.1 Introduction

The research on female labor force participation has a long tradition. Triggered by the growing labor supply of women in the second half of the last century[1], a large strand of theoretical and empirical research on this new issue has arisen. Nevertheless, the early literature on female participation in the labor market was largely concentrated on gender gaps in monetary constraints as well as social norms of working in order to explain differences between the participation decisions of men and women. Based on this literature, we already know a lot about why women keep on having lower participation rates and wage elasticities than men and why these variables started converging in the past decades.[2] However, differences in the participation probabilities between women can only be explained by differences and trends in monetary constraints as well as social norms of working to a limited. While traditional economic models largely attribute these unexplained differences in decision outcomes to idiosyncratic shocks or unobserved constraints and opportunities, modern behavioral economic and applied microeconomic approaches started to investigate these differences with respect to unobserved, inherent beliefs and preferences. A growing literature is thus interested in investigating the psychological black box behind female labor supply decisions as well as, in a more general sense, individual decision making on the labor market altogether.

This paper contributes to the literature by investigating the role of a specific personality trait, which has already been found to have important explanatory power for decision making especially on the labor market: an individual's perception of control, also called locus of control (LOC). LOC can be characterized as a *"generalized attitude, belief, or expectancy regarding the nature of the causal relationship between one's own behavior and its consequences"* (Rotter, 1966) and describes whether individuals believe in the effects of their own efforts and abilities on their lives' outcomes. While individuals with an internal LOC (internals) believe that their own efforts and abilities are rewarded in their future life's outcome, individuals with an external LOC (externals) attribute life's outcomes mainly to luck, chance, fate or other people. LOC has already been shown to have important effects on economic behavior and decision making on the labor market[3]. Nevertheless, to the best of my knowledge, no paper exists yet which investigates the effect of LOC on female participation decisions in a comprehensive and detailed manner.

Based on an one-period model of discrete labor supply decisions, which allows for heterogeneity in preferences through the incorporation of a vector of personal attributes into the individual utility function, the theoretical considerations mainly discuss an effect of LOC on participation probabilities via a difference in the direct marginal utility from participating.

[1] See Killingsworth and Heckman (1986), Blau and Kahn (2007), Costa (2000), Goldin (1990) and Mincer (1985) for comprehensive overviews over the trends in female labor force participation during the 20th century.

[2] See Section 2.2 for a detailed overview over the literature.

[3] See Cobb-Clark (2015) for detailed discussion of the concept as well as an overview of the literature on LOC in labor economics.

In line with a similar argumentation in Cobb-Clark *et al.* (2014) for the effect of LOC on health investments, women might not only derive utility from the consumption level generated through participation but also from the act of generating this consumption level itself. Based on the psychological literature which makes an important connection between locus of control and independence considerations, this direct non-monetary gain from participation is expected to be higher for internal women. Internals put a greater weight on the status of being active in the labor market as they not only derive utility from the consumption level as an outcome of participation but also from the fact that they themselves had control over the generation of it. Additionally, potential alternative mechanisms, which can largely be formalized by differences in the individual budget constraints, are discussed. LOC might, for example, have an effect on beliefs about returns to investments, such as parental investments, job search and investments into the future career, but also on objective differences in the opportunities and constraints, e.g. wages or family income.

Therefore, in the empirical part of the paper, I estimate the direct relationship between LOC and current labor force participation of a woman in a reduced form approach. The estimations are conducted using the extensive information available from the Socio-Economic Panel (SOEP, 2017), a large representative longitudinal household panel from Germany. The SOEP not only includes detailed socio-economic information but also surveys individuals' locus of control on a regular basis. Using this data, I estimate the average marginal effects of a woman's LOC on her probability of participating in the labor force using a random effects logit estimation conditional on standard socio-economic determinants of participation. In this context, labor force participation is defined as a general availability for market production and thus concentrates on the behavioral implications of LOC on labor supply decisions. I find a significant positive relationship between having an internal LOC and being available to the labor market. A subgroup analysis reveals that while a strong relationship can be observed for cohabiting women and women with (young) children, the effect for childless women is distinctly lower or even close to zero, depending on family status. This indicates a crucial heterogeneity with respect to underlying monetary incentives to work which is additionally supported by findings on similar heterogeneity with respect to the available family income in the household. In addition, a second heterogeneity analysis displays that the estimated effects are also sensitive with respect to the underlying social norms of working as measured by regional differences as well as cohort and time trends. Furthermore, additional analysis reveals important differences between labor market participation decisions at the intensive and the extensive margin. While the positive effect on labor market availability also translates into higher labor market activity, as measured by employment probabilities, for all considered sub-groups. The increased activity can only be observed at the extensive margin. For those women who do work, LOC has no significant effect on working hours.

Consequently, it is the objective of the paper at hand to contribute to the existing literature on the important economic implications of locus of control and to provide additional empirical evidence on the drivers of women's labor supply decisions. Hence, the paper's main contribution is to add valuable insides to the approach of getting closer to the bottom of the psychological black box behind female labor force participation decisions by explicitly considering labor supply decisions as the result of a woman's inherent personality traits. In so doing, the paper has two important findings. On the one hand, the paper empirically identifies important behavioral implications of locus of control in women's participation decisions above and beyond traditional monetary incentives and disincentives. On the other hand, especially the heterogeneity analysis is able to give an idea of how strongly a woman's preference-driven labor supply decision is limited by budget constraints as well as exogenously determined social norms. The second finding is likely to hold true in general, independent from the considerations about locus of control in the present analysis.

The outline of the paper is as follows. Section 2.2 gives a comprehensive overview over the related literature. Section 2.3 summarizes the theoretical idea of the paper and, based on this, proposes hypotheses for the empirical analysis. In the empirical part of the paper, Section 2.4 describes the data and the estimation strategy and Section 2.5 presents an overview over the results of the main estimation, the heterogeneity analysis as well as the additional results for alternative outcome variables. Section 2.6 summarizes a number of tests for the robustness of the results. Section 2.7 concludes the paper.

2.2 Literature Review

The paper at hand significantly adds to two important bodies of economic literature which are on the one hand the literature on female labor force participation, including determinants of participation, overall time trends and gender differences, and on the other hand the growing literature on the economic importance of personality traits in general and the personality trait locus of control in specific.

The early literature on female labor force participation largely focused on the explanation of decreasing gender participation gaps as well as the positive long-term trends in female labor force participation. Inspired by the work of Mincer (1962), especially the increase in the average wage rate of women was in the center of attention for a long time (see e.g. Mincer, 1985; Smith and Ward, 1985; Blau and Kahn, 2007; Juhn and Murphy, 1997). Large parts of this decrease in wage inequalities were explained by increasing returns to human capital for women (see e.g. Autor *et al.*, 2008; Blau, 1998; McGrattan and Rogerson, 2008). Although the wage rates of women and men did converge and the female wage elasticity fell over time (Blau and Kahn, 2007; Costa, 2000; Heim, 2007), the participation rate and the average number of working hours per week of married women in particular are still considerably lower than that of their partners and the gap in wage elasticities is also still clearly recognizable

(Evers *et al.*, 2008; Blau and Kahn, 2017). Based on the growing theoretical considerations of joint family labor supply, empirical studies additionally found a strong and stable response of female labor supply to changes in their partners' wages, whereas no such responsiveness can be identified for men (Ashenfelter and Heckman, 1974; Lundberg, 1988; Devereux, 2004). In addition to considerations about wage and cross-wage responses of female labor supply, the conventional theoretical models were largely focused on overall declines in fertility rates through, for example, the improvement of fertility control (Goldin and Katz, 2002; Bailey, 2006), the improvement of household technologies (Greenwood *et al.*, 2005), the rise of the tertiary sector (Cortes and Pan, 2018; Weinberg, 2000; Oppenheimer, 1970) and a generally increased economic demand (Angrist, 2002; Carodso and Morin, 2018) in order to explain the observed positive trends (see e.g. Costa, 2000; Blau and Kahn, 2017; Mincer, 1985; Smith and Ward, 1985).

Over the years, multiple new strands of research have evolved, which to a large extent have focused on alternative monetary factors behind (the lack of) female labor force participation such as institutional barriers and public policy (e.g. tax incentives, transfer withdrawal rates and childcare provision) (see e.g. Blundell and MaCurdy, 1999; Hausman, 1980; Eissa and Liebman, 1996) as well as costs of participation in general (e.g. transportation and childcare) (see e.g. Cogan, 1980; Gronau, 1973; Angrist and Evans, 1998). However, as summarized by Blau and Kahn (2007, 2017), all these conventional economic studies were not able to fully explain the observed trends in female labor force participation and the remaining gender gaps. Consequently, the economic research started to consider non-economic determinants of the developments, especially by discussing and empirically analyzing the role of social norms and gender role attitudes as important non-pecuniary factors (see e.g. Bertrand, 2010; Fortin, 2015; Goldin, 2006; Reimers, 1985; Costa, 2000; Carodso and Morin, 2018). This literature is in line with the sociological literature on the transformation of traditional gender roles over time (Cotter *et al.*, 2011; Ross *et al.*, 1983). As stated by Goldin (2006), one key aspect of the *"quiet revolution of women's employment"* since the 1970s is the increasing importance of work as a key aspect for a woman's social identity. These considerations are based on the economic and sociological research on the importance of social purpose as well as economic identity and status as non-pecuniary incentives of labor force participation (Jahoda, 1981; Akerlof and Kranton, 2000).[4] Traditional gender roles and the associated gender differences in the acceptance of home production as an alternative to market production (see e.g. Killingsworth and Heckman, 1986) are crucial drivers of differences in the importance of these social norms between men and women (see e.g. Bertrand *et al.*, 2015; Knabe *et al.*, 2016; Charles *et al.*, 2018). Thus, non-pecuniary incentives to work through prevalent social norms are likely to be crucial especially when explaining gender differences in labor force

[4] The main message of this literature is straightforward: individual identity influences economic outcomes since deviating from socially desirable behavior is costly for the individual (see e.g Clark, 2003; Schöb, 2013; Hetschko *et al.*, 2014).

participation as well as long-term trends in female participation decisions. Nevertheless, a lot of unexplained heterogeneity in the decision making between women remains. Consequently, especially the most recent literature started to investigate the role of inherent personal attributes for female decision making on the labor market. Wichert and Pohlmeier (2010) find that e.g. the Big-Five personality traits play a significant role in explaining women's labor supply. Although they also consider indirect effects through wage-differences, they conclude that the traits conscientiousness, extraversion, neuroticism and openness have a crucial direct behavioral effect on participation decisions.

A second important strand of literature this paper contributes to is the emerging research linking individuals' personality to their behavior on the labor market. Especially locus of control has already been shown to have a tremendous positive effect on "desirable" behavior and decision making on the labor market in such areas as human capital investment (Coleman and DeLeire, 2003), job search effort (McGee and McGee, 2016; Caliendo *et al.*, 2015b), occupational attainment (Heywood *et al.*, 2017; Cobb-Clark and Tan, 2011), entrepreneurial activity (Caliendo *et al.*, 2014) and labor market mobility (Caliendo *et al.*, 2015a). Nevertheless, literature that directly relates female labor force participation to locus of control is scarce. Most prominently, Heckman *et al.* (2006) find a significant positive effect of a combined measure of locus of control and self-esteem on the individual probability of being employed at age 30 for the sample of young individuals from the NLSY79. They show that this relationship is much more pronounced for females. In a more recent study, Berger and Haywood (2016) analyze the effect of locus of control on mothers's return to employment after parental leave. Using German survey data, they find that women with an internal locus of control return to employment more quickly. Based on a heterogeneity analysis with respect to the underlying flexibility in the women's occupations, they conclude that the effect is mainly driven by different subjective expectations about future career costs of maternity leave. That study is most closely related to the paper at hand. Nevertheless, it concentrates on a very specific group of women in a rather exceptional stage of life whereas this paper is intended to draw a much more general picture.

2.3 Theoretical Considerations

2.3.1 Basic Model of Female Labor Supply

In order to discuss and formalize the potential mechanisms through which locus of control might affect participation decisions of women, an one-period model of discrete labor supply decisions is considered as the baseline in the following (see e.g. McFadden, 1974; Borghans

et al., 2008; Almlund *et al.*, 2011). A woman's within-period discrete participation choice is based on the maximization of the following utility function:

$$U_i = U_i(C_i, L_i, H_i, P_i; \theta_i) \tag{2.1}$$

with U being the neoclassical utility function of woman i, which maps her preferences. U is convex, i.e. increasing in its arguments and twice differentiable ($U' > 0$ and $U'' < 0$). C_i is the consumption level and L_i is leisure. As an extension to the traditional labor supply model such as e.g. in McFadden (1974), two new arguments are added to the utility function. Firstly, H_i captures all sorts of non-market (home) production such as in specific childcare. Hence, childcare is also allowed to generate a positive utility for woman i. A woman i thus divides her total amount of time T between leisure, market- and home production. Secondly, P_i captures the "joy of working", i.e. the non-monetary benefits from working, independent from disutility of every hours worked (which is inversely captured by L). Similar to the argumentation in Cobb-Clark *et al.* (2014)[5], P_i is based on the fact that woman might not only derive direct utility from the consumption level but also from the behavior they undertake to generate it. Consequently, utility not only depends on consumption as the monetary outcome of participation but also is a function of participation itself, independent from how large its effect on consumption levels might be.

In line with Borghans *et al.* (2011) and Almlund *et al.* (2011), the marginal gains from all arguments depend on a vector of individual attributes θ_i. Hence, as opposed to the standard decision making model, the preferences are heterogeneous. Based on the expansion of the traditional choice model, which allows for non-stochastic personal attributes θ_i instead of stochastic idiosyncratic shocks (McFadden, 1974), the potential role of locus of control for explaining individual differences can be discussed. For simplification, the vector θ_i is reduced to the one parameter locus of control by holding all other attributes constant. For intuition, two groups of women can be defined: Internal women (I) with the vector θ_I and external women (E) with the vector θ_E. In line with its later construction, we assume θ_i to increase with internality, i.e. $\theta_I > \theta_E$.

Woman i chooses the labor force status, which maximizes her utility and fulfills her budget constraint

$$y_i + \tilde{w}_i(T - L_i - H_i) \quad \geq \quad C_i + p_h(T - H_i) \tag{2.2}$$

The non-labor income y_i summarizes all sources of income which are not generated by the woman participating in the labor market such as partner's earnings and capital income. w_i is the potential wage rate of the woman if she decides to supply a positive amount of labor

[5] Cobb-Clark *et al.* (2014) argue that investments into health can have direct effects on individual utility independent from their outcomes simply through the act of investing which might generate positive utility by itself.

to the market, i.e. $T - L_i - H_i > 0$, with T being the endowment of time. The woman is not expected to have perfect information. Thus, $\tilde{w}_i$ captures the expectations about the potential wage rate and is a subjective measure, which again depends on individual attributes: $\tilde{w}_i(\theta_i)$. The price of the aggregated consumption good is normalized to 1 and p_h is the hourly price for externally provided childcare and housework which is not provided by the woman $(T - H_i)$. $T - H_i$ could either be provided externally or by the partner in the household. Consequently, for simplification, p_h is assumed to capture the costs of external housework or childcare as well as the hourly wage of the husband.
Woman i chooses the optimal labor force status LF_i^* which maximizes her utility with the choice set B_i being either participating $(LF_i = 1)$ if $T - L_i - H_i > 0$ or not participating $(LF_i = 0)$ if $T - L_i - H_i = 0$ at the extensive margin:

$$LF_i^* = \underset{LF_i \in B_i}{\operatorname{argmax}}\{U_i\} \quad \text{with } B_i = \{0, 1\}. \tag{2.3}$$

As the focus of this paper is to analyze the behavioral aspects of labor force participation, while leaving the demand side aside, I concentrate on labor force availability as opposed to actual employment, in the following. This reduces the risk of biased results due to omitted returns in employment probability in the empirical part.[6] In line with the ILO definition of "labor force", a woman is thus assumed to participate in the labor market if she is either already employed or self-employed or if she is unemployed and intends to participate by indicating that she is searching for a job (see International Labour Organization, 2018). Thus, LF_i also equals one if the woman does not work but is available to the market by searching for a job. In this simplification, given a certain expected market wage w_i, no assumptions on labor market conditions and frictions are necessary, as LF_i^* only depends on the woman's individual decision making processes and not on her exogenous probability of finding a job, except indirectly through her expectations captured by $\tilde{w}_i$. In line with this, conditional transfer payments which are paid in response to, for example, job search efforts, such as unemployment insurance to woman i, are also captured by w_i.[7] Thus, w_i can be labeled as the "earnings from participation".

[6] Nevertheless, also the employment probability at the extensive and intensive margin is considered as an alternative outcome variable in a later stage.

[7] This assumption might be overly rigorous as welfare fraud is an ongoing political issue and important topic of the workfare literature. Individuals can abuse social benefits by pretending to search although they might not be willing to work anyways. Discussing this issue in detail is beyond the scope of this paper. Nevertheless, active search and availability to work will be an essential part of the definition of labor force participation in the empirical analysis instead of just relying on reported labor force status in order to identify truly involuntary unemployment. The interested reader may refer to the political literature which discusses welfare fraud as an important topic. (see e.g. Roosma *et al.*, 2016) Additionally, see e.g. Kreiner and Tranæs (2005) for a discussion of the role of workfare for the separation of voluntary and involuntary unemployment in general.

2.3.2 Locus of Control and Potential Mechanisms

Based on this underlying model and the definition of locus of control, multiple hypotheses can be formed about the relationship between LOC (θ_i) and female labor force participation. These hypotheses will then guide the empirical analysis in Sections 2.4 and 2.5.

Mechanism 1 - Preferences The first potential channel suggests that LOC might affect a woman's preferences for the different components of the utility function and thus the marginal utility she derives from participation. Nevertheless, two conflicting hypotheses can be proposed based on this mechanism. In the framework outlined above, woman might not only derive utility from the consumption level but also directly from participation in the labor market as the behavior they undertake in order to generate this level, i.e. P_i as a non-monetary incentive for participation. In line with the argumentation in Cobb-Clark *et al.* (2014), internal women are likely to have a higher preference for being active in the labor market than external women. Thus, they derive more additional direct utility from participation than do externals

$$\frac{\partial^2 U_i}{\partial P_i \partial \theta_i} > 0. \tag{2.4}$$

As a psychological reasoning for this, it can be argued that internals prefer to directly affect their life's outcomes and thus be independent from external forces. They are more resistant to external influences and make more independent judgements. (Lefcourt, 1982) The role of independence considerations for locus of control has already been discussed especially in the context of early childhood skill formation in the psychological literature. (see e.g. Wichern and Nowicki, 1976; Hill, 2011) Thus, consumption which is generated based on self-earned income is valued higher than consumption based on external income such as partner's earnings or unconditional social transfers. Based on these theoretical considerations, internal woman are ex ante expected to be more likely to participate if P_i is increasing with participation.

As opposed to this, in the presence of children in the household, internal women might consider the effect of their own actions on their children more carefully than external women. This is in line with the findings by Lekfuangfu *et al.* (2018) on the strong effect of maternal LOC on attitudes towards parental style as well as actual parental time investments. Thus internal mothers might have stronger preferences for home production as they gain higher utility from every unit of H_i:

$$\frac{\partial^2 U_i}{\partial H_i \partial \theta_i} > 0. \tag{2.5}$$

If a mother assumes that her efforts in child-rearing has important positive effects on her children's outcomes, she is also more likely to stay at home with young children as opposed to putting them into childcare. This channel would ex ante be associated with a negative effect

of LOC on participation probabilities in the presence of children in the household as H_i is decreasing with participation.

Mechanism 2 - Budget Constraints The second proposed channel suggests that LOC might directly affect a woman's subjective and objective budget constraints. Firstly, women might differ with respect to their expected monetary returns to participation driven by differences in the subjective expectations about returns to investments. The expected monetary returns to participation are higher for internal individuals as they believe in the direct causality between their own efforts and life's outcomes. Internal women, for example, have higher subjective job-offer arrival rates (Caliendo *et al.*, 2015b), higher appreciation of future career costs of non-working (Berger and Haywood, 2016) and higher subjective wage rates. Hence, they expect higher earnings from participation, i.e. $\partial\tilde{w}_i/\partial\theta > 0$, and thus gain higher utility from availability for market production as their budget constraints allows for higher returns to participation in expected consumption levels $\tilde{C}_i$:

$$\frac{\partial U_i}{\partial \theta_i} = \frac{\partial U_i}{\partial \tilde{C}_i} \times \frac{\partial \tilde{C}_i}{\partial \theta_i} > 0. \tag{2.6}$$

Nevertheless, the difference between internal and external women could also be driven by differences in the objective monetary returns to locus of control and thus indirectly via different constraints rather than through direct behavioral effects. One potential explanation for this may be positive demand-side responses to an internal LOC, i.e. higher realized wage rates (see e.g. Heineck and Anger, 2010) which are correctly anticipated by women and thus incorporated into the decision-making independent from the subjective beliefs discussed above. Additionally, internal women have been found to select occupations that are less open for flexible employment paths, i.e. leaving and returning to employment, such as science, engineering or related professions (Cobb-Clark and Tan, 2011). These occupations are likely to be associated with higher future career costs of non-participation and thus higher disincentives for home production through reduced future wages and employment probabilities. Thus not only $\tilde{w}_i$ but also w_i itself depends on θ_i.

Secondly, LOC might also be correlated with the partners' earnings driven by assortative mating or mating probabilities in general. Lundberg (2012) for example shows that personality traits have an important effect on the formation and stability of marriage. In the present case, it may be the case that internal women tend to marry men with higher or lower earnings or even tend to be less or more likely to marry at all, which would again affect their own participation probabilities through y_i. It is also possible that assortative mating is important with respect to the personality of the partner. Women with an internal LOC might be more likely to mate with men with an internal LOC, which again indirectly influences women's participation decisions through their partners' earnings. Hence, internal women might differ with respect to their available family income and thus have a higher or lower necessity to work

in order to achieve their desired consumption level. Nevertheless, these indirect effects of LOC on participation probabilities via different budget constraints can be largely controlled for in the empirical analysis in Section 2.6.1.

2.3.3 Effect Heterogeneity

The influence of personality on participation via preferences and beliefs can be assumed to crucially depend on the overall size of underlying incentives. If monetary and non-monetary incentives for market or home production are very high, the power of personality to affect participation probabilities may be comparably low. Thus, the estimated effects are expected to be highly heterogeneous with respect to the overall size of e.g. y_i and H_i, which are determined by the existence of partners and children in the household as well as the amount of non-labor income such as partners earnings. Partners earnings and thus the level of family income have already been found to be negatively associated with female participation probabilities (Ashenfelter and Heckman, 1974; Lundberg, 1988; Devereux, 2004). A single woman without any other external income sources ($y_i = 0$) thus has to fulfill all her basic consumption needs with her own market earnings or conditional transfer payments. If $w_i - p_h$ is assumed to be non-negative[8], the woman has to generate a certain level of $w_i(T - L_i - H_i)$ and thus $(T - L_i - H_i) > 0$. If we assume no welfare fraud taking place, this woman will certainly participate in the labor market since basic welfare benefits follow the workfare principle and are thus conditional on participation.[9] As this considerably constraints her free choice between market and home production, the power of θ_i to affect LF_i^* is expected to be very low.[10]

Also the existence of children might interact with the effect of LOC on participation probabilities because the presence of children is likely to significantly increase the monetary and non-monetary utility from home production H_i and has thus been found to significantly reduce labor force participation of women (see. e.g. Angrist and Evans, 1998; Bronars and Grogger, 1994). If no children are present in the household, the utility from home production may be too low and home production might thus be a less attractive outside option from market production. LOC is therefore expected to affect the participation probabilities of women with (non-adult) children more strongly. On the other hand, the existence of (young) children could also lead to very high non-monetary incentives for home production, depending on a woman's parenting preferences, age of the children and the quality and costs of childcare options, and thus again reduce the power of LOC.

[8] For the specific case of Germany, this assumption is reasonable as childcare costs are relatively low.

[9] Adults who receive social transfer payments in Germany are in general required to be available for any reasonable employment if they are employable (§7 SGB II). Unemployment insurance payments are thus directly bound to an active job search requirement.

[10] This consideration can be slightly softened for single mothers with young children in Germany. In German law, employment is, amongst others, not "reasonable" if this employment would, for example, endanger the upbringing of children. As is regulated in §10 SGB II, this applies to children under the age of 3. Thus for these women, unemployment insurance payments can be assumed to be captured by y_i as long as no other sources of income exist.

As a third heterogeneity consideration, the population-wide or sub-group specific amount of non-monetary utility from participation, captured by $\partial U_i/\partial P_i$, is likely to be important. If the part of the marginal effect which reflects social norms for working, is already very high, even women who individually gain lower marginal utility from participation (i.e. external women) still have a high probability of participating as the marginal utility from participation is already considerably high:

$$\left(\frac{\partial U_i}{\partial P_i}\right)_1 > \left(\frac{\partial U_i}{\partial P_i}\right)_2 \tag{2.7}$$

$$\left(\frac{\partial^2 U_i}{\partial P_i \partial \theta_i}\right)_1 < \left(\frac{\partial^2 U_i}{\partial P_i \partial \theta_i}\right)_2 \tag{2.8}$$

with (1) being the group of women which is exposed to high social norms for working and (2) being the group of women which is exposed to low social norms of working (see e.g. Jahoda, 1981; Clark, 2003; Hetschko *et al.*, 2014). This is also in line with the idea that, for example, for men the social norms of "earning your own living" or "being the breadwinner" are expected to be very high in general independent from their LOC (see e.g. Killingsworth and Heckman, 1986; Bertrand *et al.*, 2015; Knabe *et al.*, 2016; Charles *et al.*, 2018). The same might be true for groups of women who are subject to very high social norms of working. For them the harm from staying at home exceeds the gains from participation independent from their personal attributes θ_i. Within the analysed population in the empirical part of the paper, variation in social norms is expected to be captured by regional differences as well as cohort and time trends.

2.4 Data and Empirical Identification

On the basis of these theoretical considerations, the goal of this paper is to empirically analyze the role of locus of control in explaining women's current labor force participation. This is done by using data from the German Socio-Economic Panel (SOEP, 2017). The SOEP is an annual representative household panel that follows a general-purpose approach. It has been studying about 22,000 individuals living in 12,000 households in Germany since 1984. Personal questionnaires are completed by all individuals aged 18 or older. For more information on the SOEP see Goebel *et al.* (2018). It contains a measurement of locus of control over multiple waves, rich information on current labor-market outcomes and family status as well as the opportunity to connect women to regional information as well as their partners' characteristics if they are surveyed in the same household. The data is restricted to the waves 2000-2015 due to the measurement of locus of control as well as the availability of the regional information.

Sample Restriction For the sample restriction process, I only keep women in the traditional working age, which is defined as 25 to 65 years. As another restriction, all women who are still in school, academic or vocational education, already in (early) retirement or in military service are dropped. Additionally, only women who live in single-adult or in couple households with or without children are kept. All women in multi-generation households or other unknown household combinations are dropped. Finally, only women for whom it is possible to observe all the relevant socio-economic control variables are kept. This leaves 57,308 observations for 7,724 women over 15 years in the full sample of all women.

Table 2.A.1 in the Appendix gives an overview of the descriptive statistics of the full sample (column 1). In addition to the full sample of all women, the descriptive statistics are also reported for three (non-exclusive) subsamples: (1) all women with biological children under 16[11] (2) non-cohabiting women, i.e. single women or women with a partner outside their household, and (3) cohabiting women, i.e. married women or women with a partner in the same household.

2.4.1 Labor Force Participation

Labor force participation (LF) is measured as a binary indicator that indicates a woman's availability to the labor market. The focus of this paper is to analyze the behavioral aspects of labor supply. Thus, labor force participation does not describe a woman's true labor force status but her willingness to participate in market production. Concentrating on the availability to the job market rather than on the actual employment status allows the demand side to be theoretically and empirically neglected and the risk of selection effects via differences in employment probabilities to be reduced.

In line with the ILO definition of labor force participation, a woman is counted as being in the labor force if she is either employed or self-employed or if she is registered unemployed or non-working (not registered unemployed) but intends to work and is searching for a job (see International Labour Organization, 2018)[12]. The robustness of the results with respect to the definition of the dependent variable is tested and discussed in detail in a sensitivity check in Section 2.5.4. For this purpose an indicator for labor force activity (employment), indicators for differences at the intensive margin (e.g. full-time employment) as well as cumulative time in the labor force during the core working age are used as alternative outcome variables.

[11] The sample includes all women independent of their family status. The information on the children is generated on the basis of the individual birth history of the women available from the SOEP. The sample thus includes all women who have children under the age of 16 (independent of whether they live in the same household).

[12] Registered unemployed and non-working women are recoded on the basis of the information available on intention to work, active search and ability to start working from the personal questionnaire. Registered unemployed women who indicate that they were not actively searching for work in the last 4 weeks are coded to "not participating" while women who were originally coded as "not working" but indicate that they actively searched for a job, have the unconstrained intention to work and are ready to immediately start working are coded to "participating".

Table 2.1: Descriptive Statistics - Labor Force Status

	(1) All	(2) Children under 16	(3) Not Cohabiting	(4) Cohabiting
	mean	*mean*	*mean*	*mean*
Labor Force Status				
Employed	0.70	0.62	0.77	0.69
Full-Time ($\geq$ 35 hours)	0.54	0.35	0.68	0.50
Part-Time (15 - 34 hours)	0.37	0.51	0.28	0.40
Marginal ($<$ 15 hours)	0.09	0.15	0.04	0.11
Unemployed	0.07	0.07	0.12	0.05
Self-Employed	0.06	0.06	0.07	0.06
Not-Working	0.13	0.15	0.03	0.16
Maternity Leave	0.04	0.10	0.01	0.04
Outcome: Labor Force Participation	0.82	0.74	0.93	0.79
Observations	57,308	22,219	11,267	46,041
Individuals	7,724	3,617	2,305	6,565

Source: SOEP, waves 2000 - 2016, version 33, own calculations.
Note: Full descriptive statistics can be found in Table 2.A.1 in the Appendix.

Table 2.1 gives an overview of the current labor force status of the women. In the full sample of all women (column 1), 70% are employed, 6% are self-employed, 7% are unemployed and in total 17% indicate that they are not working or on maternity leave. If, in addition to these raw shares, the information on active job search, intention to work and availability to start working are also considered, a labor force participation rate of 82% results. Due to a lower share of employed women and a higher share of women who indicate that they are not working in the subsample of cohabiting women (column 4) as compared to the subsample of non-cohabiting women (column 3), the labor force participation is distinctively lower for the former (79% compared to 93%). The same holds true for all women with children under the age of 16 (column 2).

2.4.2 Locus of Control

Locus of Control is surveyed within the SOEP in the years 1999, 2005, 2010 and 2015. Based on a scale developed by Nolte *et al.* (1997), respondents were asked how closely a series of 10 statements characterizes their views about the extent to which they influence what happens in life. A four-point Likert scale ranging from 1 ('applies fully') to 4 ('does not apply') was used in 1999, while in 2005, 2010 and 2015 responses were measured on a seven-point Likert scale ranging from 1 ('disagree completely') to 7 ('agree completely'). A list of the items can be found in Table 2.2.

Table 2.2: Components of Locus of Control (not imputed)

No	Item	*mean*	*SD*
Q:	The following statements apply to different attitudes towards life and the future. To what degree do you personally agree with the following statements? *Scale: 1 (Disagree completely) - 7 (Agree completely)*		
I1:	How my life goes depends on me	5.46	(1.39)
I2:	Compared to other people, I have not achieved what I deserve (-)	3.15	(1.76)
I3:	What a person achieves in life is above all a question of fate or luck (-)	3.53	(1.63)
I4:	If a person is socially [...] active, she can have an effect on social conditions	3.71	(1.58)
I5:	I have the experience that others have a controlling influence over my life (-)	3.11	(1.66)
I6:	One has to work hard in order to succeed	5.91	(1.14)
I7:	If I run up against difficulties in life, I often doubt my own abilities (-)	3.49	(1.66)
I8:	The opportunities that I have in life are determined by the social conditions (-)	4.54	(1.43)
I9:	Inborn abilities are more important than any efforts one can make	4.78	(1.31)
I10:	I have little control over the things that happen in my life (-)	2.63	(1.47)
	Observations	14,214[a]	

Source: SOEP, waves 1999, 2005, 2010 and 2015, version 33, doi:10.5684/soep.v33.
Notes: Items marked with a (-) are reversed prior to factor analysis.
[a] In this table, the item means and SD are computed for the observation waves 1999,2005, 2010 and 2015 only. Imputed values are not included.

In order to harmonize the scales, the responses from 1999 are reversed and "stretched".[13] Afterwards, an exploratory factor analysis is conducted separately by year in order to investigate the way these items load onto latent factors.[14] Items 1 and 6 clearly load onto the first factor – which is interpreted as internal LOC –, while items 2, 3, 5, 7, 8 and 10 clearly load onto the second factor – interpreted as external LOC. Items 4 and 9 are not included in the following due to ambiguous loading[15]. Additionally, similar to the reasoning in Specht *et al.* (2013), excluding those two items improves the internal consistency and scale reliability of the resulting factor as Cronbach's alpha (Cronbach, 1951) increases from 0.62 to 0.67.

In line with the previous literature (see e.g. Piatek and Pinger, 2016), a two-step procedure is used in order to create a continuous and unidimensional LOC factor. First, the scores for items 2, 3, 5, 7, 8 and 10 are reversed such that all eight items are increasing in internality.[16] Second, confirmatory factor analysis is used to extract a single factor for each year. This has the advantage that it avoids simply weighting each item equally, as averaging would do, and instead allows the data to determine how each item is weighted in the overall index. Simple averaging of all items would risk measurement error and attenuation bias (Piatek and Pinger,

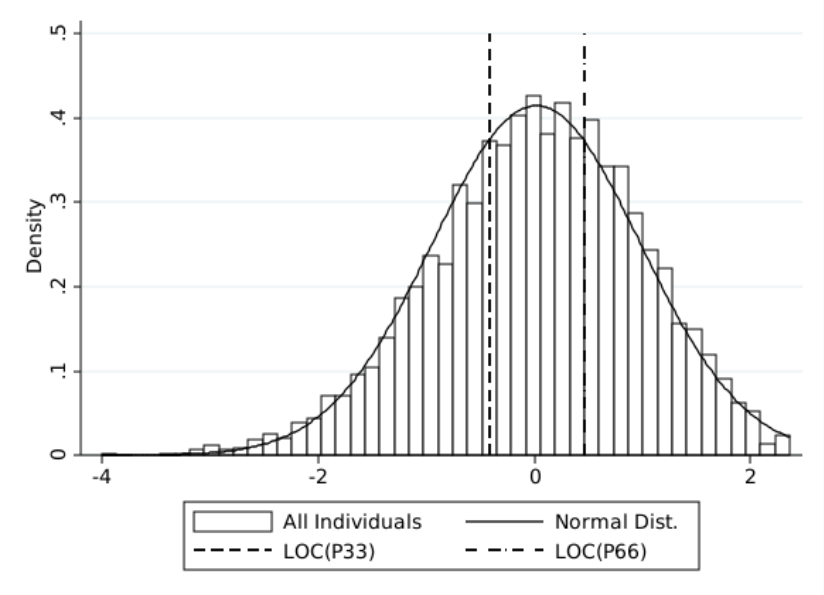
[13] In line with Specht *et al.* (2013), this process preserves the relative differences between individuals. The process results in values of 1, 3, 5 or 7 such that a '1' on the 1999 four-point scale, for example, becomes a '7' on the 2005-2015 seven-point scales.

[14] The detailed results from the factor analysis are available upon request. A loading plot pooled for all years can be found in Figure 2.A.1 in the Appendix but the pattern of factor loadings is similar in all years.

[15] Item 4 does not clearly load onto either of the two factors and is discarded. Item 9 loads onto the internal factor but an intuitive attribution based on the item's wording would point more in the direction of an external item.

[16] These items are marked with a (-) in Table 2.2.

Figure 2.1: Distribution of Locus of Control



Source: SOEP, waves 1999, 2005, 2010 and 2015, version 33, doi:10.5684/soep.v33, own illustration.

2016).[17] The resulting factor is increasing in internal LOC and its distribution is shown in Figure 2.1.

On the basis of the generated and imputed continuous LOC factor variable, I create a categorical variable that splits the continuous LOC in three terciles, in order to identify non-linear relationships. These cutoffs are also illustrated in Figure 2.1.

There is evidence that LOC is relatively stable for the working-age population (see e.g. Preuss and Hennecke, 2018; Cobb-Clark and Schurer, 2013). Nevertheless, in order to minimize concerns about potential reverse causality, I ensure that the LOC factor is always included as a pre-market rather than a contemporaneous or post-market measure, i.e. always obtained prior to t, in line with the argumentation in Piatek and Pinger (2016). Hence, LOC is imputed forwards lagged by at least one year. That is, LFP in 2000 - 2005 depends on the 1999 locus of control, LFP in 2006 - 2010 depends on the 2005 locus of control and LFP in 2011 - 2015 depends on the 2010 locus of control.[18]

[17] The sensitivity checks in Section 2.6.2 include a re-estimation of the results using this simple index. The results do only marginally differ.

[18] Based on the findings in Preuss and Hennecke (2018), this procedure does not prevent a bias due to a temporary measurement error in LOC during periods of unemployment. In line with what is proposed in Preuss and Hennecke (2018), I thus additionally attempt to correct the LOC measurement by using the LOC which has been observed during the closest employment spell of these women. The results of these estimations are presented and discussed in Section 2.6.2. As an additional sensitivity check also the average over all available LOC observations is used as an explanatory variable.

2.4.3 Estimation Strategy

For the main empirical analysis, I employ a reduced-form approach to estimate the association between a woman's propensity to be available to the labor force and her last LOC:

$$P(LF_{it} = 1) = P(\beta_1 + \beta_2 loc_{it-n} + \beta_3 X_{it} + \beta_4 P_i + \beta_5 R_{it} + \beta_6 T + \epsilon_{it} > 0), \qquad (2.9)$$

where LF_{it} is the indicator for labor force participation of woman i at time t and loc_{it-n} is the locus of control of woman i in the last LOC interview prior to t, i.e. n interviews prior to t with $n = \{1, \ldots, 6\}$. In order to identify potential non-linearities in the relationship, the analysis is repeated with a categorical variable that indicates in which tercile of the LOC distribution a woman is classified. The vector X_{it} contains an extensive list of demographic information (age, religion, region of residence, school and vocational degree, subjective health) and family characteristics (partner status, number of children, indicators for children in certain age ranges). Additionally, family income is approximated by subtracting the reported labor net income as well as individual unemployment insurance payments from the reported net household income. The variable is thus assumed to capture all earnings which are not generated through labor force participation of woman i. It includes partners earnings as well as all other regular income sources such as unconditional social transfer payments.[19] It is contains as a dichotomous indicator for a low, medium or high family income. Additionally, averaged and standardized personality and preferences measures P_i (Big Five personality traits and risk aversion) are included as controls.[20]

The vector R_{it} contains a list of regional information corresponding to the local planning region ("Raumordnungsregion")[21] in which the woman lives at time t. The information is available from the INKAR data provided by the German Federal Institute for Research on Building, Urban Affairs and Spatial (BBSR, 2015). The vector includes an indicator for East-Germany, the unemployment rate, the gross value added, the population density, the median full-time income of women, the share of children in public childcare for the age groups 1-3 years and 3-6 years as well as the share of full-time (> 7 hours per day) childcare in the respective age-groups. Additionally, the median costs for full-time childcare per child are approximated

[19] Only earnings-related unemployment insurance benefit, the so called "Arbeitslosengeld I" are subtracted. Unemployment benefits as part of the basic income scheme for needy jobseeker, the so called "Arbeitslosengeld II", cannot be distinguished from other social assistance payments within the household and is thus still captured within the variable although it still follows the principle of activity. See e.g. Konle-Seidl *et al.* (2007) for an overview over the unemployment benefits system in Germany.

[20] In a sensitivity check in Section 2.6, I additionally investigate the role of characteristics of the employment type in the current or last employment spell in order to assess the importance of selection in certain industries or occupational types. Nevertheless, these variables are assumed to be "bad controls" in the estimation due to the high risk of endogeneity of those variables in the model (see Angrist and Pischke, 2008, for more information). See a more detailed discussion of the problem in Section 2.6.

[21] The households residential location is classified into one of 96 separate regions using the SOEP geocodes. Although these regions do not correspond to official local government areas, they are the basis for the federal German government's regional planning. In particular, they capture urban centers (along with their associated catchment areas) and are defined on the basis of commuting flows (see BBSR, 2015).

based on the information available from all SOEP respondents in the respective region.[22] See Table 2.A.1 for the full list of controls. Finally, the vector T contains year fixed effects.

Equation 2.9 is estimated using a random effects logit model. Hence, ϵ_{it} captures the idiosyncratic error as well as the random individual-specific effect which is assumed to be uncorrelated with the explanatory variables. The use of random effects is necessary as it considers the panel structure of the data and takes care of serial correlation of the error term ϵ_{it} across time for a given individual i. The results presented in Section 2.5 are the average marginal effects on the probability of a positive outcome assuming that the random effect is zero.[23]

As already discussed in the theoretical considerations, the estimated direct relationship is very likely to be non-linear with respect to monetary and non-monetary incentives as important decision constraints. Consequently, heterogeneity with respect to the family status (i.e. existence of a partner and children in the household) and the level of family income as well as with respect to underlying differences in social norms of working (i.e. region of living, cohort indicators and period indicators) is considered in a second step. Since not only β_2, i.e. the marginal effect of *loc*, is regarded to be non-linear, the heterogeneity is examined using fully separated models for the different subgroups SG_{it}:

$$P(LF_{it} = 1|SG_{it}) = P(\beta_1 + \beta_2 loc_{it-1} + \beta_3 X_{it} + \beta_4 P_i + \beta_5 R_{it} + \beta_6 T + \epsilon_{it}|SG_{it}). \quad (2.10)$$

In order to prevent problems with selection into these sub-groups depending on locus of control, the LOC is generated, standardized and cut into terciles for each sub-group separately such that women are always only compared to women in the same sub-group. All results from the heterogeneity analysis were also replicated using interactions instead of the fully separated models. The results differ only marginally and can be obtained from the author upon request.

2.5 Results

2.5.1 Descriptive Evidence

Table 2.3 gives first descriptive evidence for the relationship between LOC and labor force status and participation of the women in the sample. The shares of all labor force statuses as well as the dependent variable labor force participation are given separately for all three terciles of LOC. It can be seen that due to a higher share of employed and self-employed

[22] In the years 2002, 2005, 2007, 2011, 2013 and 2015 households report detailed information on the childcare costs per child in the household as well as information on the type and amount of childcare. Based on all households with a child in full-time public childcare (> 7 hours) in the region, the median over all reported amounts is calculated.

[23] In order to assess the sensitivity of the results with respect to the choice of method, the main results were replicated using a simple logit model with clustered standard errors and a random effects linear probability model. The estimated results of both alternatives only marginally differ from the random effects logit estimation in effect size and are available from the author upon request.

Table 2.3: Descriptive Statistics by LOC - Labor Force Status

	Low LOC $[LOC_{min}, LOC_{P33}]$	**Medium LOC** $(LOC_{P33}, LOC_{P66}]$	**High LOC** $(LOC_{P66}, LOC_{max}]$
Labor Force Status			
Employed	0.67	0.72	0.72
Full-Time Employed	0.53	0.53	0.55
Part-Time Employed	0.47	0.47	0.45
Unemployed	0.10	0.06	0.04
Self-Employed	0.04	0.06	0.08
Not-Working	0.16	0.12	0.11
Maternity Leave	0.03	0.04	0.04
Outcome: Labor Force Participation	0.78	0.82	0.84
Observations	18,912	18,917	19,479

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.

women and a lower share of non-working women for the highest tercile, the overall share of labor force participation is higher for women with a high LOC than for those with a low LOC. Nevertheless, this descriptive relationship is very likely to be driven by a long list of socio-demographic characteristics that are associated with a higher participation probability and a higher LOC, such as education, age and family status (e.g. number and age of children).

2.5.2 Main Results

Table 2.4 presents an overview of the estimated average marginal effects of the continuous LOC variable and the LOC categories medium ($(LOC_{P33}, LOC_{P66}]$) and high ($(LOC_{P66}, LOC_{max}]$) on labor force participation, with a low LOC ($[LOC_{min}, LOC_{P33}]$) being the reference category. All the estimations are for the full estimation samples of all women, while gradually including more and more sets of control variables. In addition, column 1 of Table 2.A.2 in the Appendix provides an overview of the estimated marginal effects for all the control variables analogous to the estimation using the continuous LOC variable (columns 7 in Table 2.4).

In line with the descriptive evidence in the previous subsection, the results of the raw difference, only controlled for year fixed effects, indicate that on average women with a high or medium LOC are ceteris paribus more likely to participate in the labor force (see column 2). Furthermore, the continuous LOC factor is significantly positive, indicating an increasing probability of participation with increasing values of LOC (column 1). Including additional control variables indicates that the raw gap was biased downwards by omitted-variable bias especially through family characteristics. The effect size increases considerably between between columns 3 (4) and 5 (6), when partner status, number of children, age of children and family income are included. Thus, in the full specification (columns 7 and 8), the average marginal effect is significantly positive and even more pronounced as in the raw difference. Having a medium or a high LOC increases the probability of being in the labor force by, on

Table 2.4: Main Results - Average Marginal Effects (Sample: All)

	Outcome Variable: Labor Force Participation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LOC Factor (cont.)	0.006***		0.004***		0.010***		0.008***	
	(0.001)		(0.001)		(0.002)		(0.002)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.009***		0.009***		0.015***		0.013***
		(0.002)		(0.003)		(0.003)		(0.003)
$(LOC_{P66}, LOC_{max}]$		0.012***		0.009***		0.019***		0.014***
		(0.002)		(0.003)		(0.004)		(0.004)
Observations	57,308	57,308	57,308	57,308	57,308	57,308	57,308	57,308
Year Fixed-Effects	✓	✓	✓	✓	✓	✓	✓	✓
Regional Controls	✓	✓	✓	✓	✓	✓	✓	✓
Socio-Demographic Controls			✓	✓	✓	✓	✓	✓
Family Controls					✓	✓	✓	✓
Personality Controls							✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Full estimation results for the specification in column (7) can be found in Table 2.A.2 in the Appendix.

average, 1.3 and 1.4 percentage points compared to having a low LOC. When comparing this effect to the mean non-participation rate in the full sample of 18 percent (see Table 2.1), this amounts to a 7.8 percent decrease in the probability of staying at home. Increasing the LOC by approximately one standard deviation, increases the probability of participation by 0.8 percentage points (column 7).

When comparing the marginal effects of a medium and a high LOC, a non-linearity in the effect of LOC on the participation probability becomes apparent. While a medium LOC is associated with an increased probability of participation, this effect flattens out afterwards. Women with a very high LOC are not significantly more likely to participate than women with a medium LOC. In line with the one-dimensionality of the LOC scale, the findings indicate that the effect is mainly driven by a negative impact of being strongly external, rather than a positive impact of being strongly internal.

2.5.3 Heterogeneity Analysis

Family Status and Children Using these main results for the full sample, Table 2.5 presents the results for the sub-samples based on family status and existence of biological children under the age of 16. All estimations include the full specification and, for comparison, columns 1 of Table 2.5 give the results for the full sample (equivalent to column 8 in Table 2.4).[24] Fully separated estimations are reported. These subgroup analyses correspond

[24] In addition, Table 2.A.2 in the Appendix gives an overview of the estimated marginal effects for all the control variables corresponding to the estimations in columns 1, 3 and 7. The full results for all other models can be obtained from the author upon request.

Table 2.5: Heterogeneity Analysis (Marginal Effects): Family Status and Children

	All			**Non-Cohabiting**			**Cohabiting**		
	Children under 16			Children under 16			Children under 16		
	All	No	Yes	All	No	Yes	All	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)									
$(LOC_{P33}, LOC_{P66}]$	0.013***	0.007***	0.022***	0.001	-0.004	0.028**	0.015***	0.013***	0.016**
	(0.003)	(0.002)	(0.007)	(0.004)	(0.002)	(0.014)	(0.004)	(0.003)	(0.008)
$(LOC_{P66}, LOC_{max}]$	0.014***	0.009***	0.022***	0.004	-0.001	0.023	0.015***	0.014***	0.015*
	(0.004)	(0.003)	(0.008)	(0.004)	(0.002)	(0.015)	(0.005)	(0.004)	(0.009)
Observations	57,308	35,089	22,219	11,267	8,404	2,863	46,041	26,685	19,356
LF = 0	10,526	4,688	5,838	822	380	442	9,704	4,308	5,396
LF = 1	46,782	30,401	16,381	10,445	8,024	2,421	36,337	22,377	13,960
	(82%)	(87%)	(74%)	(93%)	(95%)	(85%)	(79%)	(84%)	(72%)
All Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Full estimation results for the specifications in columns (1), (4) and (7) can be found in Table 2.A.2 in the Appendix.

to the supposed heterogeneity of the effect of LOC on participation probabilities with respect to underlying monetary and non-monetary incentives and disincentives to work driven by the existence of partners and children in the household.

Looking at the estimated average marginal effects for the separate groups, we can see that the effect is in large part driven by cohabiting women and women with children under 16 in general. Cohabiting women with a medium/high LOC are, on average, ceteris paribus 1.5 percentage points more likely to be in the labor force than cohabiting women with a low LOC (column 7). For non-cohabiting women, the effect is insignificant and close to zero[25]. However, separating the two groups by the children-indicator shows that this difference is in large part driven by the existence of children in both groups. In the subgroup of women with children under 16, the effect for non-cohabiting women, i.e. single mothers, (column 6) is even higher than the effect for cohabiting women with children (column 9). Single mothers with a medium LOC are, on average, ceteris paribus 2.8 percentage points more likely to be in the labor force than single mothers with a low LOC. A high LOC increases the probability of being in the labor force by 2.3 percentage points although not significant due to the small sample size. In the sub-sample of women without children, the effect is still positive and significant for cohabiting women (column 8) but zero for non-cohabiting women (column 5). Table 2.A.3 in the Appendix also provides analog estimation results for mothers depending on the age of their children. Women with pre-school children exhibit the largest effects (column 3 of Table 2.A.3). Having a high LOC on average decreases a woman's non-participation

[25] Similar insignificant close-to-zero effects can be found for men using the same sample specifications. The results can be obtained from the author upon request.

probability by 6 percentage points (18%) if she has pre-school children. For mothers of babies especially the effect of a high LOC is close to zero. As opposed to this, the effect for women with "adult" children is still significant and positive.

All these results support the theoretical idea that the effect of LOC on participation probabilities strongly interacts with underlying incentives and disincentives to work. If the monetary incentives for market production, such as in the case of single women without children, already considerably exceed the decision threshold, personality and preferences have no power to affect the participation decision. The same holds true for very high monetary disincentives to work, e.g. in the sample of women with children under the age of one, i.e. women who are largely eligible for generous parental leave payments. Thus, very high monetary incentives and disincentives to work act in the same direction, i.e. they reduce the power of LOC.

Also in line with the theoretical considerations, a certain level of monetary and non-monetary incentives for home production impose the necessary scope of decision making that is important in order for LOC to change the decision towards working. This is, for example, given through the social purpose the existence of (young) children adds to home production. Being independent of monetary incentives, only women who at least consider home production as an outside option to market production are assumed to be affected by their LOC. If no young children are present in the household, the utility from home production seems to be very low and home production is thus a less attractive outside option to market production – such as in the case of cohabiting women without children (column 8 in Table 2.5) or with school-age and adult children (columns 4 and 5 in Table 2.A.3). Evaluating the theoretical ideas on the consequences of very high non-monetary incentives for home production, e.g. through the importance of care provided by the mother as opposed to outside-household childcare, is less straightforward as this is highly heterogeneous with respect to individual preferences about childcare and the quality of alternative childcare options. The theoretical idea that an internal LOC might be associated with a lower participation probability for mothers due to considerations about their own influence on children's outcomes cannot be found in the results for participation probabilities in general. The effect of LOC on participation probabilities is robustly positive over all the subgroups considered. Merely the more pronounced non-linearity in the effects for mothers (especially single mothers) might be driven by this alternative mechanism.[26]

Family Income As a more direct measure of differences in the monetary incentives to work, the approximated family income can be used for the definition of the sub-groups. Family income is intended to capture all non-labor income y available to a women in the household,

[26] The sensitivity check in Section 2.5.4 reveals remarkable differences at the intensive margin, showing that mothers with a medium LOC are actually less likely to work full-time. This is clear evidence for considerations about own children's outcomes being at play when women decide about working hours.

Table 2.6: Heterogeneity Analysis: Family Income

	Family Income		
	Low	Medium	High
	(1)	(2)	(3)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.010***	0.012*	0.010
	(0.003)	(0.006)	(0.007)
$(LOC_{P66}, LOC_{max}]$	0.010***	0.013*	0.016**
	(0.003)	(0.007)	(0.007)
Observations	18,835	18,783	19,690
LF = 0	1,486	3,822	5,218
LF = 1	17,349 (92%)	14,961 (80%)	14,472 (74%)
All Controls	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

e.g. partners labor income and unconditional social transfers. For a detailed description on the generation of this variable, see Section 2.4.3. The three sub-groups are generated based on whether the family income is low, medium or high with the terciles being the cut-offs. Table 2.6 presents the results of the heterogeneity analysis. All estimations include the full specification using the dichotomous LOC measure (equivalent to column 8 in Table 2.4). The sample of all women is used and fully separated estimations are reported.

The observation numbers and reported shares in the bottom panel of Table 2.6 show that women with a high family income are, in line with expectations, distinctly less likely to participate in the labor force (74%) as compared to women with a low family income (92%). The sub-group analysis reveals indication for the expected heterogeneity. While the marginal effects of a medium LOC are rather consistent, the effect of a high LOC seems to increase with increasing family income. While internal women with a low family income are on average only one percentage point more likely to participate, the participation probability for internal women with a high family income is on average 1.6 percentage points higher than for external women.

Social Norms of Working In additional to budget constraints, a woman's decision making might also be constraint by prevailing social norms for working. If social norms for working are rather high (such as e.g. for men), decision making based on inherent preferences is considerably constraint. Table 2.7 presents the results of this heterogeneity analysis. All estimations include the full specification using the dichotomous LOC measure (equivalent to column 8 in Table 2.4). The sample of all women is used and fully separated estimations are reported.

Firstly, heterogeneity can be expected with respect to differences between the eastern and western parts of Germany. Due to the long-term socialist political influence in the former

Table 2.7: Heterogeneity Analysis: Social Working Norms (Sample: All)

	Region		**Cohort**[1]			**Observation Period**		
	West	East	Early	Middle	Late	'99 - '05	'05 - '10	'10 - '15
			<'58	*'58-'66*	*>'66*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$	0.016***	0.007*	0.020***	0.008***	0.005	0.031***	0.008	0.005
	(0.005)	(0.004)	(0.007)	(0.003)	(0.006)	(0.011)	(0.007)	(0.006)
$(LOC_{P66}, LOC_{max}]$	0.021***	0.003	0.025***	0.004	0.016***	0.020*	0.017***	0.016***
	(0.005)	(0.004)	(0.008)	(0.003)	(0.006)	(0.012)	(0.007)	(0.006)
Observations	41,770	15,538	18,507	17,747	21049	16,619	23,000	21,049
LF = 0	8,871	1,655	3,959	2,327	4,240	3,782	4,131	2,613
LF = 1	32,899	13,883	14,553	15,420	16,809	12,837	18,869	15,076
	(79%)	(90%)	(79 %)	(87%)	(80%)	(77%)	(82%)	(85%)
All Controls	✓	✓	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[1] Cohort Cutoffs: Early - born before 1958, Middle - born 1958-1966, Late - born after 1966.

GDR, the east of Germany has a longer tradition of women's participation in the labor force.[27] Consequently, direct marginal utility from participation $\partial U_i \backslash \partial P_i$ is expected to be higher for eastern German women and the absolute effect of LOC on participation probabilities is likely to be lower. The observation numbers (columns 1 and 2 in bottom panel of Table 2.7) support this assumption. The participation probability is with 79% distinctly lower in the west of Germany than in the east of Germany (90%). The upper panel of columns 1 and 2 of Table 2.7 presents the marginal effects based on the subgroups of women living in the east and west of Germany. The results reveal that the significant positive marginal effect of a medium and a high LOC is distinctively larger for women in the west of Germany. The effect of a high LOC is not significantly different from zero for women in the east of Germany. The identified heterogeneity might nevertheless also be driven by differences in the availability of childcare between the east and west of Germany. Nevertheless, shares of children in public childcare for different age-groups are included as controls in all estimation models. Additionally, a heterogeneity analysis with respect to those shares is less clear and gives no indication for heterogeneity in this respect.[28]

Based on the continuous decrease in the importance of traditional gender roles over time in almost all modern Western societies (see e.g. Goldin, 2006), women in later cohorts are assumed to be more affected by a generalized social pressure to be economically independent

[27] The socialist system was characterized by a strong emphasis on the dual-earner/state-carer system of family labor supply, i.e. an extremely high levels of female labor force participation in combination with an extensive system-level organization of family-support structures and child care (see e.g. Braun *et al.*, 1994; Rosenfeld *et al.*, 2004).
[28] Results are available from the author upon request.

from external forces than women of earlier cohorts (Heim, 2007). For the former, $\partial U_i \backslash \partial P_i$ can be assumed to be higher than for the latter. They might therefore have a higher participation probability independent from LOC as their utility from participation is higher. Thus, columns 3 to 5 of Table 2.7 present the marginal effects based on the estimations in which the marginal effect of LOC is calculated fully separated for different cohorts. The cutoffs for the manifestations of the birth cohort indicator "early", "middle" and "late" were generated based on the terciles of year of birth in the full estimation sample, i.e. $P(33) = 1958$ and $P(66) = 1966$, in order to obtain groups of approximately similar size. The results indicate a strong heterogeneity of the effect with respect to cohort. The distinct marginal effects of a medium and high LOC on participation probabilities can only be observed for women from the early cohorts, i.e. born before 1958 (column 3). The effect is distinctly lower for both the women in the medium as well as in the latest cohorts (columns 4 and 5).

As the heterogeneity with respect to cohort might also be driven by age effects, the heterogeneity was repeated by making use of the longitudinal nature of the data. The observations are thus cut into three groups based on the observation periods 1999-2005, 2005-2010 and 2010-2015. The cutoffs are informed by the LOC observation waves 2005 and 2010. The results reported in columns 6 to 8 support the identified heterogeneity in the previous sub-group analysis. The effect size seems to decrease over time, supporting the idea of increasing social norms for working.

2.5.4 Labor Force Activity, Working Hours and Lifetime Participation

The behavioral implications of LOC on labor force availability has been the center of attention in the theoretical considerations as well as the main part of the empirical analysis. Nevertheless, it is interesting to investigate whether those static behavioral effects actually translate into higher employment probabilities and higher average lifetime participation, as these are the variable with the desired positive macro- and microeconomic consequences in the long run. If a higher probability of being available to the market for internal women does not translate into higher employment probabilities, the positive economic implications of LOC are limited by other unobserved factors such as e.g. market conditions and frictions.

In order to assess the generalizability of the results with respect to the choices made about the participation indicator as described in Section 2.4.1, three major components of the dependent variable are investigated: 1) the concentration on labor force availability instead of labor force activity, 2) the restriction to the extensive margin as well as 3) the focus on an one-period discrete choice rather than a lifetime perspective on labor force participation.

Table 2.8 starts by giving the results of the investigations with respect to the first two points. As a first step, the dependent variable is adjusted such that it only captures labor force activity instead of availability. Thus, the indicator is one if a women is actually employed and zero if she is unemployed or not-working independent from whether she intents to work by

Table 2.8: Additional Results (Marginal Effects): Outcome Variable - Labor Force Activity

				Sample: All Employed					
	Employment			No Marginal			Full-Time		
	All	Kids <16	Cohab.	All	Kids <16	Cohab.	All	Kids <16	Cohab.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)									
$(LOC_{P33}, LOC_{P66}]$	0.028***	0.031***	0.031***	-0.001	-0.004	-0.000	0.010	-0.004	-0.007
	(0.005)	(0.010)	(0.006)	(0.002)	(0.006)	(0.002)	(0.010)	(0.015)	(0.011)
$(LOC_{P66}, LOC_{max}]$	0.030***	0.032***	0.031***	0.001	-0.005	0.002	0.013	-0.006	-0.003
	(0.005)	(0.011)	(0.007)	(0.002)	(0.006)	(0.002)	(0.011)	(0.017)	(0.012)
Observations[1]	53,907	20,936	43,376	40,225	13,874	31,548	40,225	13,874	31,548
LF = 0	13,682	7,062	11,828	3,736	2,021	3,350	18,606	9,078	15,828
LF = 1	40,225	13,874	31,548	36,489	11,853	28,198	21,619	4,796	15,720
	(75%)	(66%)	(73%)	(91%)	(85%)	(89%)	(54%)	(35%)	(50%)
All Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[1] Self-employed are dropped from the sample as working hours cannot be determined for them.

searching for a job.[29] This alternative definition was neglected in the main part of the empirical analysis as it captures unobserved returns to locus of control with respect to employment probabilities and therefore does not concentrate on the behavioral aspects of labor force participation.

Columns 1 to 3 of Table 2.8 give the results of this new indicator while still concentrating on the extensive margin. The results indicate that the behavioral changes are fully translated into higher employment probabilities. The effects are considerably stronger than in the main estimations. This is likely due to unobserved returns to LOC in employment probabilities. Having a high LOC thus on average increases the probability of being employed by 3 percentage points for the full sample (column 1). This is also true for the sub-samples of mothers (column 2) and cohabited women (column 3).

In addition to this, columns 4 to 9 give the estimated marginal effects of LOC on participation indicators at the intensive margin. For the sub-sample of all women who are employed ($LF = 1$ in columns 1 to 3), the outcome variable in columns 4 to 6 indicates whether the woman is employed at least part-time and consequently not marginally employed, defined as 15 contracted working hours per week or more. Additionally, the outcome variable in columns 7 to 9 indicates whether a woman is full-time employed, defined by at least 35 contracted working hours per week. While LOC positively affects labor force availability as well as participation at the extensive margin, no effects can be identified at the intensive margin. This

[29] In this specification, self-employed women are set to missing as it is not possible to identify working hours for these women in the further steps. Nevertheless, the results for the indicator at the extensive margin do not strongly differ if self-employed women are included. Results are available upon request.

Table 2.9: Additional Results (Marginal Effects): Outcome Variable - Aggregated Participation (Sample: 55+)

	Sample: Cross-Section - Women 55+					
	Years in Labor Force 25-55y			Years Employed 25-55y		
	All	Kids[2]	Cohab.	All	Kids[2]	Cohab.
	(1)	(2)	(3)	(4)	(5)	(6)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)[1]						
$(LOC_{P33}, LOC_{P66}]$	0.180	0.216	0.251	0.759**	0.782**	0.622
	(0.351)	(0.376)	(0.402)	(0.360)	(0.384)	(0.406)
$(LOC_{P66}, LOC_{max}]$	0.835**	0.701*	0.716*	1.531***	1.399***	1.158***
	(0.372)	(0.396)	(0.423)	(0.381)	(0.404)	(0.428)
Observations	3,255	2,909	2,640	3,255	2,909	2,640
All Controls	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[1] Locus of Control is calculated based on the average over all available LOC observations.
[2] The sub-group consists of all women with any biological children at time t independent from their age.

is largely in line with the theoretical idea of LOC mainly affecting the non-monetary gains from participation independent from the amount of working hours.

Additionally, the lifetime perspective needs to be considered in order to get an idea about whether this static relationship actually translates into differences for the whole working life. Thus, in the additional results presented in Table 2.9, the accumulated years in the labor force as well as in employment between the age of 25 and 55 are the outcome variables of interest. Using the detailed biographical information available for every SOEP participant, the aggregated time in the labor force is calculated by adding the years a woman spent in employment or registered unemployment during those 30 years.[30] As no biographical information is available on the job-search behavior, I have to rely on the reported labor force status in order to identify labor force participation here. As job-search is likely to be an important determinant of true willingness to participate, it has to be taken into account that this is thus only a rough measure of participation. The cross-sectional estimation sample consists of one observation per woman in the first available year in the sample in the age of 55 or later. Only women who have valid information on their labor force status during at least 25 of those 30 years are considered. The explanatory variable is a measure for the average locus of control over all available observations. This is largely in line with the sensitivity reported in Section 2.6.2 and is intended to reduce the problem of regression attenuation in the present

[30] A women is assumed to spend a full year in a certain labor force status if she only reports one spell during a certain year. If she reports multiple spells during one year, she is assumed to have spend an equal share of the year in either spell and consequently the value (1\number of spells) is added to the counter.

case as LOC at time t might be measured long after the relevant decision periods.[31] The effects are estimated using a simple linear regression model.

The results presented in Table 2.9 indicate a significant positive effect of LOC on lifetime labor force availability and activity during those 30 years. Women with a high average LOC spend on average approximately 0.8 years more in the labor force (column 1) and 1.5 more years in employment (column 4) during this time.[32] The effect is relatively homogeneous with respect to the marital status as well as indicator for children (of any age) at time t.

2.6 Sensitivity Analysis

2.6.1 Controlling for Endogenous Budget Constraints

As has been discussed in the theoretical considerations, differences in the participation probabilities between internal and external women might also be driven by omitted differences in the objective budget constraints rather than the proposed differences in preferences or beliefs. Thus, controlling for them is necessary to identify the direct behavioral effect of LOC on participation decisions instead of the indirect effects through differences in opportunities and constraints such as occupational selection, wage differences and assortative mating. The results from this additional estimations can be found in Table 2.A.4 in the Appendix.

Controlling for those potentially endogenous variables is, however, less straightforward than expected. Simply including the information on the current or last job would leave us with a large multicollinearity problem caused by the characteristics themselves, but also by their availability in general. The information on employment characteristics (occupation and wage) has to be imputed from the last employment or self-employment spell if a woman is not (self-)employed at the moment. Nevertheless, it is not possible to observe any information on employment for a lot of women if they were either never employed or at least never employed during their time in the SOEP[33]. This is, by definition, more often the case in the group of women who do not participate in the labor force at the moment. When controlling for the employment information, the indicator for non-availability of the information would thus be a "bad control", in line with the arguments by Angrist and Pischke (2008), as it is highly multicollinear with the labor force participation indicator. Not only are external women more likely to be observed outside the labor force at the moment, but they are also more likely never to be

[31] Although smaller in magnitude, the effects are still significant if instead the imputed LOC at time t is used. Results are available upon request.

[32] The lower estimates for labor force participation are likely to be driven by the rough measure of participation here as no information on job search is available and also all other alternative labor force statuses such as retirement and education influence the overall number of years. As internal women are observed to spend less years in registered unemployment, this drives the estimated results downwards.

[33] The problem can be slightly weakened by using information on the very first employment of women given in the biography questionnaire, which is answered by every SOEP-respondent. If no information on occupational type or industry is available during the time in in the SOEP, this information is used. Nevertheless, no wage information is available from the biographical data.

observed in the labor force, and the indicator could just as well be a dependent variable in the estimation model. To disentangle the endogeneity problem from the true effects of controlling for occupational characteristics and wages, column (1) of Table 2.A.4 starts by reducing the observation sample to the women who are observed in occupation during their time in the SOEP at least once. In line with expectations, the estimated effects for the LOC drop if the sample is reduced, indicating an endogeneity problem in the observability of information. The estimated effects from this reduced sample are taken as the new baseline in the following in order to eradicate parts of the bad controls problem. Nevertheless, it has to be taken into account that also the occupational information themselves are potentially endogenous. Therefore, the presented results have to be interpreted with care.

Occupational Choice Column 2 of Table 2.A.4 starts by adding potentially omitted information on the industry and occupational type of women in their current or last job as measures of occupational selection based on locus of control. Nevertheless, if the last occupational type and the last industry classification are included, the estimated effects for LOC do not change. Hence, this can be taken as an indication of no severe bias through omitted occupational information in the main estimations. An effect of LOC on participation probabilities via occupational selection and differences in the expected future costs of non-participation can thus be rejected.

Wage Differences Secondly, it is necessary to consider wage returns to LOC as potential omitted drivers of the identified effects. In order to investigate the importance of this channel, a proxy for the expected wage from working has to be included as a control variable. Column 3 of Table 2.A.4 presents the results of the sensitivity check in which additionally to the occupational information also the net labor income of the last observed working spell is included as a proxy for the expected wage of a potential future employment. If the reduced sample is used and the last labor net income is included as a control variable, this does not further change the estimated coefficient for LOC. The coefficient of the last net labor income has the expected positive sign. Omitted variable bias can therefore be rejected with respect to a demand-side response to LOC via higher expected wages, too.

Assortative Mating - Partner's Wage and Personality As a third set of variables that might explain parts of the estimated relationship, information on a woman's partner has to be controlled for. Fortunately, the SOEP makes it possible to merge cohabiting women with their partners. Thus, columns 4 to 6 of Table 2.A.4 present the results of the sensitivity check in which the partner's current net labor income as well as the continuous LOC factor of the partner is included as an additional control variable for cohabiting women. In line with the procedure in the paragraphs above, the results of the baseline estimation for the reduced sample of all women for whom it is possible to merge the partner's wages and locus

of control are presented in column 4 of Table 2.A.4. As can be seen in columns 5 and 6, the results do not change if partner's net income and LOC are included as control variables, indicating that the results of the main estimation are not severely biased by assortative mating. As partners labor income is already captured by the indicator for family income in the main estimations, the effect of partners income is, although significantly negative, rather small. Partners LOC has no significant effect at all. Admittedly, remaining concerns about various other possible interdependencies between partners such as those with respect to gender attitudes cannot be completely ruled out. Internal women might, for instance, be more likely to mate with men who have more tolerant gender attitudes and are thus more likely to participate. These characteristics are assumed to be largely captured by the partner's locus of control and earnings. The fact that the positive relationship between LOC and participation decisions can also be identified for single mothers, however, provides some reassurance that assortative mating does not drive large parts of the estimated relationship.

2.6.2 Locus of Control Construction and Imputation

As a second important set of sensitivity checks, the construction and imputation of the locus of control as explanatory variable is tested. Table 2.A.5 in the Appendix presents the results of three alternative forms of construction and imputation of the LOC factor: (a) a simple index, (b) an average over all observations as well as (c) an imputation from the closest employment spell. All checks are conducted using the full estimation sample and the full set of control variables. The baseline results are replicated in columns 1 and 2. Columns 3 and 4 provide the re-estimated effects when varying the construction of the LOC factor in the observation years. Instead of using the results of the factor analysis, the "simple index" assumes equal weights of all 8 items used, and a simple average over all item responses is calculated. The results differ only marginally from those in the baseline.

Secondly, the timing of the LOC measurement and thus the imputation approach is tested. As Preuss and Hennecke (2018) pointed out, there is a considerable risk of reverse causality or attenuation bias due to temporary measurement errors in the LOC. Using the same data from the SOEP, they found a significant negative short-run effect of exogenous job-loss on LOC for individuals who are still unemployed during the LOC interview. Based on an extensive sensitivity check, they conclude that this is likely to be driven by temporary state-dependent reporting in the LOC for unemployed individuals even though LOC can be assumed to be stable in the long-run. Due to the fact that employed and non-employed individuals are pooled in my estimation sample, there might be a risk of biased results due to a measurement bias in LOC, which would, by definition, be greater in the group of non-participating women due to a higher share of non-employed individuals in this group[34]. In order to circumvent this measure-

[34] While in the group of participating women potentially only some of the women, i.e. those who are unemployed, might have a state-bias in their observed LOC, the share is expected to be greater in the group of non-participating women as 100% of women in this sample might be affected by such a state-bias.

ment problem, two alternative approaches are implemented. Firstly, instead of the forward imputed LOC, a variable which averaged all available LOC observations of an individual between 1999 and 2015 is used as the explanatory variable. This approach is likely to reduce the attenuation bias in the LOC due to temporary measurement errors to a minimum. The results of this alternative estimation are presented in columns 5 and 6 of Table 2.A.5. In line with expectations about biases due to regression attenuation, the estimated effects increase considerably if this alternative LOC measure is used. This gives indication for a downward bias due to measurement error in the LOC variable in the main estimations. Although these new results might again be biased upwards by reverse causality, the main estimations are likely to only be lower bounds of the true effects.

Using the average LOC nevertheless does not solve problems with reverse causality if the measurement error is selective as women who are not employed in t have a higher probability to also be not employed in the periods before and after t. Those women thus always report a lower LOC due to their non-activity on the labor market. Therefore, additionally the LOC observation during the closest employment or self-employment spell to t is used. The two conditions for imputing the LOC observation from a period $t + x$ or $t - x$ into t are that (a) LOC has to be observed in that year and (b) the woman is observed to be employed or self-employed in that year.[35] Nevertheless, this approach has one main caveat: by imputing from the closest employment spell, all women who are never observed in (self-)employment are lost. This is largely in line with the problems discussed in the first paragraph of the sensitivity section. Never being observed in (self-)employment is highly endogenous to the model in line with the argumentation above. Columns 7 and 8 of Table 2.A.5 thus check the effect of the LOC variable in the baseline model, using only the sample of women for which the LOC variable from the closest employment is observed. As expected, although still positive and significant, the estimated effect is now considerably smaller, indicating a problem with endogeneity in the observability of employment spells. Based on this reduced sample, columns 9 and 10 present the results for the alternative approach of imputation for the LOC factor. When using the reduced sample, the alternative LOC variable actually increases the estimated effects. Thus, if the main estimations are at risk of being biased, this is likely to be a bias towards zero as opposed to an overestimation of effects.

2.7 Conclusion

How do women make decision about their labor force participation at a given point in time and what factors determine heterogeneity in participation probabilities between and within genders? This is a question economists have already been interested in for many years

[35] I also allow for backwards imputation to avoid problems with sample size. This is based on the assumption that, besides measurement bias in LOC through non-employment, non-employment has no long-term effect on LOC based on the findings in Preuss and Hennecke (2018).

of fruitful theoretical and empirical research. Nevertheless, we are still far from solving the puzzles within this long-lasting "hot topic" in labor economics. While a lot of open questions have been answered on the gender gap in labor market participation with especially wages gaps and differences in social norms for working and gender roles being key determinants of the differences between men and women, the prevalent within-gender heterogeneity still keeps economists and politicians busy. When getting to a point at which standard economic theories are unable to fully explain individual decision making with observed opportunities and constraints, traditional approaches often contribute remaining heterogeneity to stochastic idiosyncratic shocks (see e.g. McFadden, 1974; Borghans *et al.*, 2008). Nevertheless, with the advancement of modern behavioral economics and applied microeconomics in combination with an increased availability of comprehensive survey-data, economists got more and more interested in understanding the psychological black box behind unexplained individual differences.

In line with this, especially in the last couple of years, empirical economic research was increasingly informed by ideas from psychology. A huge body of empirical literature provides evidence for the importance of especially personality traits for individual decision making in various domains of the individual life-cycle. This paper contributes to this line of research in the context of female labor force participation by theoretically and empirically discussing the role of the personality trait locus of control for differences in participation probabilities between women. Due to the rich facets of the construct locus of control, it can be assumed to influence multiple components of a woman's maximisation problem when choosing the optimal labor force status. Nevertheless, in line with the existing literature, a crucial role of LOC for independence preferences and expected returns to investment decisions and thus a positive relationship with the direct marginal utility from participation but also home production through subjective monetary and non-monetary gains is expected.

Based on the theoretical considerations, a reduced form estimation of the relationship between LOC and a woman's probability of being available to the labor market is the center of the empirical analysis. Using a random effects logit model, I find that internal women, i.e. women who believe in the importance of their own efforts for life's outcomes, are on average more likely to be available to the labor force. LOC thus adds explanatory power to the participation decision above and beyond traditional socio-economic factors as well as other preferences measures. Hence, the paper significantly adds to the existing economic literature on female labor force participation as well as the important economic consequences of locus of control by suggesting and empirically identifying distinct behavioral implications of LOC in the participation decision. Hence, the paper primarily contributes to the investigation of the psychological black box behind female labor force participation and additionally broadens the knowledge on the economic importance of locus of control.

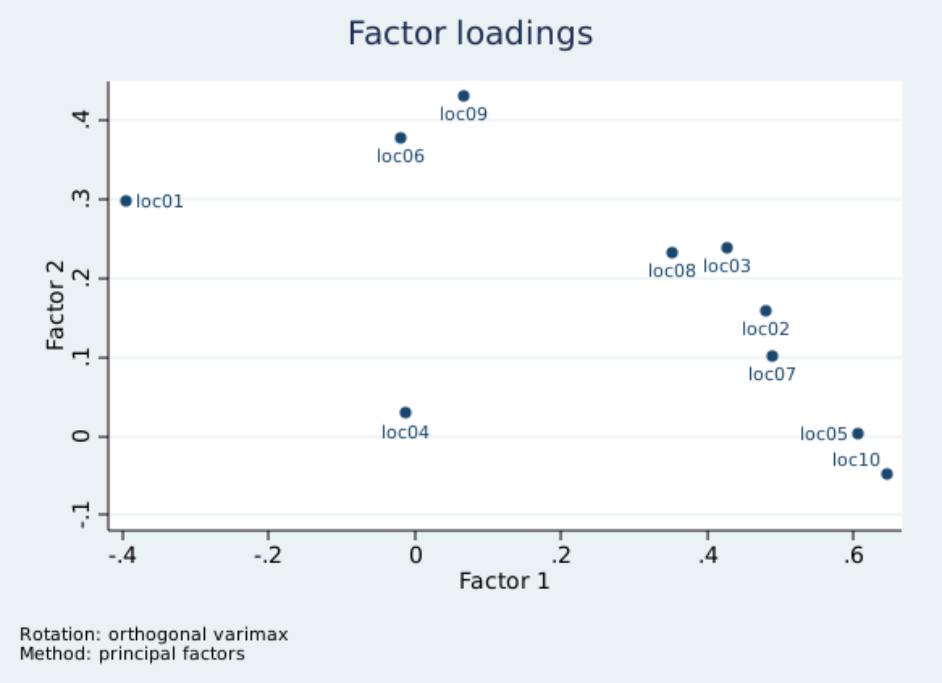
On the other hand, a heterogeneity analysis identified an interesting sensitivity of the effect with respect to given monetary constraints as well as prevalent social working norms. This suggests that inherent traits, preferences and tastes are only able to inform participation decisions if the underlying budget constraints are fulfilled and if the decision-making is not constrained by exogenous imposed social norms. It seems natural to argue that this is not a phenomenon which is specific to locus of control but very likely also translates to other measures of psychological traits and economic preferences. This neatly fits into the recent findings by Falk and Hermle (2018), who identified a similar importance of economic development for gender-inequality on the macro-level.[36]

The identified role of locus of control for a woman's decision-making process as well as the prevalent importance of exogenous constraints in the relationship has crucial implications for the widespread political discourse about low labor force participation rates of women. When discussing and evaluating political measures targeted at increasing participation rates, such as active labor market policies, quotas or childcare availability and costs, it is therefore extremely important to understand the boundaries of monetary incentives set by latent psychological characteristics. Considerations about the effectiveness of active labor market policies need to be aware of the large component in individual decision making which cannot be influenced by monetary incentives as it is based on inherent personal attributes and preferences for either participation or home production. As opposed to this, the results from the heterogeneity analysis also illustrates that preference-based decision making is massively bounded by exogenous monetary and non-monetary constraints. Reducing them would presumably raise individual welfare as individual's freedom of choice is increased. Although locus of control arguably is a trait which is of specific interest for participation considerations, due to its focus on effort and independent action, those findings are highly relevant also above and beyond this special case and the considerations are likely to be universally valid also for other traits and preferences.

[36] They identified that higher levels of economic development as well as gender equality favor the manifestation of gender differences in economic preferences under the rational that greater availability of material resources removes the universal goal of subsistence and more gender-equal access to those resources allows women and men to express preferences independently from each other.

2.A Appendix

Figure 2.A.1: Factor Loadings of the LOC Variable



Source: SOEP, waves 1999, 2005, 2010 and 2015, version 33, own illustration.

Table 2.A.1: Descriptive Statistics

	(1) All	(2) Children under 16	(3) Not Cohabiting	(4) Cohabiting
	mean	*mean*	*mean*	*mean*
Family Controls				
Family Status	.			
Single	0.13	0.08	0.68	
Partner not in HH	0.06	0.04	0.32	
Partner in HH	0.10	0.09		0.13
Married	0.70	0.78		0.87
Number of Children	1.62	2.03	1.19	1.72
Children Age Indicators				
Child under 1	0.03	0.07	0.01	0.03
Child 1 - 3	0.06	0.15	0.02	0.07
Child 3 - 7	0.13	0.33	0.07	0.14
Child 7 - 16	0.28	0.72	0.20	0.30
Family Income				
Low	0.33	0.22	0.86	0.20
Medium	0.33	0.37	0.09	0.39
High	0.34	0.42	0.05	0.42
Socio-Demographic Controls				
Age Categories				
25 - 34 Years	0.21	0.31	0.28	0.19
35 - 44 Years	0.30	0.52	0.27	0.31
45 - 54 Years	0.31	0.16	0.28	0.31
55 - 65 Years	0.18	0.00	0.17	0.19
Religion				
Non	0.32	0.27	0.37	0.31
Christian	0.64	0.67	0.61	0.65
Muslim	0.02	0.03	0.01	0.02
Other	0.02	0.03	0.01	0.02
In Bad Health	0.14	0.09	0.17	0.13
Highest School Degree				
No School Degree	0.02	0.02	0.01	0.02
Lower Secondary School	0.24	0.21	0.24	0.24
Middle School	0.40	0.42	0.37	0.41
Highschool	0.28	0.28	0.33	0.26
Other School	0.06	0.07	0.05	0.07
Highest Vocational Degree				
No Vocational Diploma	0.15	0.16	0.16	0.15
Apprenticeship	0.43	0.42	0.41	0.43
Higher Technical College	0.28	0.30	0.28	0.29
College or University Degree	0.24	0.21	0.26	0.23
Regional Controls				
East-Germany	0.27	0.24	0.29	0.27
Unemployment Rate	9.08	9.00	9.13	9.07
Gross Value Added	54.00	53.75	54.54	53.86
Population Density in 100	5.36	5.06	6.31	5.13
Median Full-Time Income Women	2551.62	2559.18	2565.92	2548.12
Public Childcare				
Children under 3y in Care	23.08	21.59	24.39	22.75
Share of Full-Time Care (<3y)	44.98	43.38	46.01	44.72
Children 3 - 6y in Care	92.31	92.15	92.33	92.30

Continued on next page...

... continued from previous page

	(1) All	(2) Children under 16	(3) Not Cohabiting	(4) Cohabiting
Share of Full-Time Care (3-6y)	37.04	35.01	38.31	36.72
Median Costs for Full-time Care	98.37	99.56	97.36	98.62
Personality Controls				
Willingness to take risk (std., avg.)	-0.20	-0.18	-0.10	-0.22
Openness (std., avg.)	0.03	0.01	0.12	0.01
Conscientiousness (std., avg.)	0.07	0.03	0.07	0.07
Extraversion (std., avg.)	0.10	0.16	0.08	0.10
Agreeableness (std., avg.)	0.18	0.20	0.13	0.19
Neuroticism (std., avg.)	0.15	0.16	0.12	0.16
Observations	57,308	22,219	11,267	46,041
Individuals	7,724	3,617	2,305	6,565

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.A.2: Main Results (Marginal Effects) - Full Results

	All	**Children under 16**	**Cohabiting**
	(1)	(2)	(3)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.013***	0.022***	0.015***
	(0.003)	(0.007)	(0.004)
$(LOC_{P66}, LOC_{max}]$	0.014***	0.022***	0.015***
	(0.004)	(0.008)	(0.005)
Family Controls			
Family Status (Ref.: Single)			
Partner not in HH	0.006	0.006	
	(0.006)	(0.019)	
Partner in HH	0.011**	0.064***	
	(0.006)	(0.016)	
Married	-0.015***	0.049***	
	(0.005)	(0.015)	
Number of Children	-0.020***	-0.034***	-0.023***
	(0.002)	(0.005)	(0.003)
Has Child under 1	-0.750***	-0.723***	-0.775***
	(0.014)	(0.014)	(0.013)
Has Child 1 - 3 Years	-0.314***	-0.308***	-0.343***
	(0.013)	(0.015)	(0.014)
Has Child 3 - 7 Years	-0.044***	-0.027***	-0.044***
	(0.005)	(0.006)	(0.005)
Has Child 7 - 16 Years	-0.005	0.027***	-0.003
	(0.003)	(0.009)	(0.004)
Family Income (Ref.: Low)			
Medium	-0.039***	-0.087***	-0.044***
	(0.003)	(0.006)	(0.003)
High	-0.085***	-0.177***	-0.096***
	(0.005)	(0.009)	(0.005)
Socio-Demographic Controls			
Age Categories (Ref.: 25 - 34 Years)			
35 - 44 Years	0.013***	0.027***	0.016***
	(0.004)	(0.008)	(0.004)
45 - 54 Years	0.007	0.038***	0.007
	(0.005)	(0.011)	(0.006)
55 - 65 Years	-0.068***	-0.002	-0.084***
	(0.008)	(0.046)	(0.010)
Religion (Ref.: No Religious Affiliation)			
Christian	0.001	0.002	-0.002
	(0.004)	(0.010)	(0.006)
Muslim	-0.053***	-0.117***	-0.066***
	(0.016)	(0.029)	(0.020)
Other	-0.034***	-0.071***	-0.059***
	(0.013)	(0.025)	(0.018)
In Bad Health	-0.025***	-0.029***	-0.026***
	(0.004)	(0.009)	(0.005)
Highest School Degree (Ref: No Degree)			
Lower Secondary School	0.018	0.071**	0.015
	(0.015)	(0.029)	(0.020)
Middle School	0.065***	0.117***	0.070***
	(0.014)	(0.033)	(0.018)
Highschool	0.067***	0.126***	0.070***
	(0.011)	(0.027)	(0.015)
Other School	0.031**	0.060**	0.033**
	(0.013)	(0.027)	(0.017)
Highest Vocational Degree (Ref.: No Vocational Diploma)			
Apprenticeship	0.032***	0.046***	0.034***
	(0.006)	(0.012)	(0.008)
Higher Technical College	0.044***	0.062***	0.047***

Continued on next page...

... continued from previous page

	(1)	(2)	(3)
	(0.005)	(0.011)	(0.007)
College or University Degree	0.053***	0.077***	0.064***
	(0.005)	(0.011)	(0.006)
Regional Controls			
East-Germany	-0.004	-0.005	0.013
	(0.020)	(0.042)	(0.023)
Unemployment Rate	-0.002**	-0.000	-0.002
	(0.001)	(0.002)	(0.001)
Gross Value Added	-0.000	-0.000	-0.000
	(0.000)	(0.001)	(0.000)
Population Density in 100 0.000	0.000	0.001	
	(0.000)	(0.001)	(0.001)
Median FT Income Women	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Public Childcare			
Children under 3y in Care	0.001**	0.001	0.001*
	(0.000)	(0.001)	(0.001)
Share of Full-Time Care (<3y)	-0.000	-0.001*	-0.000
	(0.000)	(0.000)	(0.000)
Children 3 - 6y in Care	0.002***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)
Share of Full-Time Care (3-6y)	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)
Median Costs for Full-time Care	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Personality Controls			
Willingness to Take Risk (std., avg.)	0.013***	0.007	0.019***
	(0.004)	(0.007)	(0.005)
Openness (std., avg.)	-0.004	-0.003	-0.006
	(0.003)	(0.006)	(0.004)
Conscientiousness (std., avg.)	0.021***	0.022***	0.023***
	(0.003)	(0.006)	(0.004)
Extraversion (std., avg.)	0.011***	0.016***	0.015***
	(0.003)	(0.005)	(0.003)
Agreeableness (std., avg.)	-0.005	-0.006	-0.004
	(0.003)	(0.006)	(0.004)
Neuroticism (std., avg.)	-0.013***	-0.009	-0.017***
	(0.003)	(0.006)	(0.004)
Observations	57,308	22,219	46,041
Year Fixed-Effects	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.A.3: Heterogeneity Analysis: Age of Children

	Children under 16[1]				**Adult Child**
	Baby *0-1*	Toddler *1-3*	Pre-School *3-7*	School Age *7-16*	*over 16*
	(1)	(2)	(3)	(4)	(5)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)					
$(LOC_{P33}, LOC_{P66}]$	0.021	0.018	0.041***	0.010**	0.006*
	(0.020)	(0.031)	(0.016)	(0.005)	(0.004)
$(LOC_{P66}, LOC_{max}]$	0.003	0.044	0.060***	0.011**	0.013***
	(0.019)	(0.032)	(0.017)	(0.005)	(0.004)
Observations	1,565	3,392	7,314	12,050	23,873
LF = 0	1,319	1,747	2,400	1,820	4,165
LF = 1	246 (16%)	1,645 (49%)	4,914 (67%)	10,230 (85%)	4,165 (86%)
All Controls	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[1] The groups are not mutually exclusive. Women are included if they have at least one child in the respective age-group.

Table 2.A.4: Additional Results (Marginal Effects): Confounders (Sample: All)

	Sample: Ever Employed			Sample: With Partner		
	(1)	(2)	(3)	(4)	(5)	(6)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)						
$(LOC_{P33}, LOC_{P66}]$	0.008***	0.008***	0.007***	0.017***	0.017***	0.017***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
$(LOC_{P66}, LOC_{max}]$	0.010***	0.010***	0.010***	0.016***	0.015***	0.015***
	(0.002)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
Occupational Type in Last Employment (Ref.: Blue-collar Worker)						
White-collar Worker		0.010***	0.008***			
		(0.003)	(0.003)			
Civil Servant		0.023***	0.014**			
		(0.006)	(0.007)			
Other Occupations		-0.068***	-0.064***			
		(0.023)	(0.022)			
Not Available		-0.062***	-0.062***			
		(0.023)	(0.023)			
Industry in Last Employment (Ref.: Manufacturing)						
Agriculture		-0.033**	-0.031**			
		(0.014)	(0.014)			
Mining, Quarrying, Energy, Water		0.026***	0.026**			
		(0.009)	(0.010)			
Chemicals, Pulp, Paper		0.010*	0.011*			
		(0.006)	(0.006)			
Construction		0.007	0.009			
		(0.007)	(0.008)			
Iron/Steel		-0.000	-0.001			
		(0.009)	(0.009)			
Textile/Apparel		-0.036***	-0.034**			
		(0.014)	(0.014)			
Wholesale/Retail		-0.005	-0.002			
		(0.005)	(0.005)			
Transport/Communication		0.004	0.006			
		(0.007)	(0.007)			
Public Service		0.011**	0.014***			
		(0.004)	(0.005)			
Financials/ Private Services		0.001	0.004			
		(0.005)	(0.005)			
Other		0.006	0.010*			
		(0.005)	(0.006)			
Not Available		-0.073***	-0.072***			
		(0.011)	(0.011)			
(Last) Net Labor Income (KEUR)			0.015***			
			(0.002)			
Partner						
Net Labor Income (KEUR)					0.004***	0.004***
					(0.001)	(0.001)
Partners LOC Factor (cont.)						0.000
						(0.002)
Observations	53,751	53,751	53,751	39,915	39,915	39,915
All Controls	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.A.5: Sensitivity Analysis (Marginal Effects): Locus of Control (Sample: All)

	Full Sample						Filled Employment Sample			
	Baseline		Simple Index[1]		Average[2]		Baseline		Closest Employment[3]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LOC Factor (cont.)	0.008***		0.008***		0.012***		0.004***		0.004***	
	(0.002)		(0.002)		(0.002)		(0.001)		(0.001)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)										
$(LOC_{P33}, LOC_{P66}]$		0.013***		0.012***		0.016***		0.005***		0.009***
		(0.003)		(0.003)		(0.006)		(0.002)		(0.002)
$(LOC_{P66}, LOC_{max}]$		0.014***		0.012***		0.022***		0.007***		0.010***
		(0.004)		(0.004)		(0.006)		(0.002)		(0.002)
Observations	57,308	57,308	57,308	57,308	57,308	57,308	50,129	50,129	50,129	50,129
All Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2000 - 2016, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[1] LOC Factor for each year is calculated using a simple average of all 8 items. The Index is then imputed using the same rule as in the baseline.
[2] LOC Factor is calculated for each year as in the baseline but is imputed as an average over all available LOC observations.
[3] LOC Factor is calculated for each year as in the baseline but is imputed from the closest LOC observation in which the individual was employed or self-employed and not from the last LOC observations.

CHAPTER 3

Locus of Control and Internal Migration

*This chapter is based on joint work with **Marco Caliendo**, **Deborah Cobb-Clark** and **Arne Uhlendorff**.*

3.1 Introduction

Internal migration is fundamental to the process of economic adjustment. The large-scale movement of workers in response to relative economic opportunities shapes the nature of economic disparity across geographic regions (see Blanchard *et al.*, 1992; Niebuhr *et al.*, 2012) and, in many countries, is a key driver of regional demographic change (e.g. Borjas *et al.*, 1992; Gabriel and Schmitz, 1995). Internal migration in principle reduces labor market rigidities, including structural unemployment, allowing markets to operate more efficiently. Policy makers therefore often wish to support the unemployed in migrating to stronger labor markets, while at the same time discouraging migration in response to more generous welfare benefits (e.g. De Giorgi and Pellizzari, 2009; Valletta, 2013). Incentives to promote internal migration such as lump sum grants, housing vouchers, employment and relocation services, and subsidized moving costs have been used in a variety of countries with mixed success (see Caliendo *et al.*, 2017, for a review). In particular, while incentives may lead to more internal migration, employment outcomes are not always improved as a result.[1]

Economists generally conceptualize internal migration as a fairly standard human capital investment in which individuals (and households) weigh the current costs of migration against the appropriately discounted future returns. Migration occurs whenever the expected benefits outweigh the expected costs. In other contexts, however, traditional models of this sort are increasingly giving way to models with more realistic psychological foundations. The result has been a deeper understanding of the important role that psychological traits (e.g. personality, non-cognitive skills, perceptions of control, etc.) play in most economic decisions. It is important that we begin to also incorporate the key insights of behavioral economics into models of the migration decision. Many psychologists argue, for example, that while unfavorable economic conditions may make emigration either more or less likely, the decision to stay or go rests largely on individual personality (Boneva and Frieze, 2001). In particular, Frieze and Li (2010) argue that mobility decisions are driven by individuals' desire to change their lives in ways that better satisfy their achievement, power, and affiliation motivations, while Bauernschuster *et al.* (2014) find that better educated and more risk-tolerant individuals are more likely to migrate across cultural (linguistic) regions which, the authors argue, stems from their lower psychic costs of migration. We need to know more about which psychological traits predispose certain individuals to migrate and why.

The goal of this paper is to advance the literature by incorporating locus of control into an economic model of internal migration. Locus of control can be characterized as "a generalized attitude, belief, or expectancy regarding the nature of the causal relationship between one's own behavior and its consequences" (Rotter, 1966, p.2). Those believing that life's outcomes

[1] Katz *et al.* (2001); Ludwig *et al.* (2005); Kling *et al.* (2007) and Ludwig and Kling (2007), for example, investigate the effectiveness of the Moving to Opportunity (MTO) program – introduced in the 1990s in the U.S. – and find that it successfully relocated families to better neighborhoods and partly improved their health. However, there was no significant effect on either educational or labor market outcomes.

are due to their own efforts have an internal locus of control, while those believing that outcomes are due to external factors (e.g. luck) have an external locus of control. We begin by modelling migration across domestic labor markets as the outcome of a job search process. Because they believe that search effort influences the offer arrival rate, individuals with an internal locus of control are predicted to engage in more intensive, geographically-dispersed job search – and if job search precedes migration – be more likely to migrate as a result. We then empirically test the relationship between locus of control and the propensity to migrate across regions using data from the German Socio-Economic Panel (SOEP, 2017). We find that not only do individuals with an internal locus of control express more willingness to migrate, they do in fact migrate more often. Moreover, having an internal locus of control has an effect similar in magnitude to that of key demographic and human capital characteristics such as the presence of children in the household or educational attainment.

Explicitly modeling internal migration as the result of a job search process is an important contribution to the existing migration literature. Research in labor or urban economics frequently models residential location and job search in tandem (e.g. den Berg and Gorter, 1997; van Ommeren *et al.*, 1999; Eliasson *et al.*, 2003; Lutgen and Van der Linden, 2015). Migration research, on the other hand, typically either ignores any wage uncertainty (Borjas *et al.*, 1992) or simply assumes that migration decisions occur before destination wages are realized (e.g. Harris and Todaro, 1970; Hunt, 2006; Arntz *et al.*, 2011; Kennan and Walker, 2011). Consequently, internal migration decisions depend solely on expected (actual) home vs. destination wages and any job search is simply subsumed in the aggregate employment probabilities. This lack of attention to the role of job search is surprising given that it is often disparities in unemployment rates rather than wage levels that empirically drive internal migration (e.g. Treyz *et al.*, 1993; Parikh and van Leuvensteijn, 2002). Importantly, we demonstrate that assumptions about the timing of job search have fundamental implications for the pattern of internal migration that results. Conditional on skills, internal migration in the standard model is predicted to be unidirectional with high-skilled individuals migrating to regions where their skills are more highly valued and low-skilled individuals migrating in the opposite direction (Borjas *et al.*, 1992). In contrast, modeling internal migration as the result of a job search process implies that those skills that lead to more intensive (or more productive) search in the destination labor market will lead to multidirectional internal migration, i.e. an increased propensity to migrate overall. Moreover, we adopt a more behavioral approach to modeling job search by assuming that job seekers do not know the true job offer arrival rate, but instead form subjective beliefs – related to their locus of control – about the impact of search effort on the probability of receiving a job offer. This provides a theoretical connection between locus of control and internal migration.

Our empirical findings also make an important contribution by adding weight to the emerging literature linking individuals' perceptions of control to their human capital invest-

ments through the returns that they anticipate. In particular, locus of control is related to investments in education (e.g. Wang *et al.*, 1999; Coleman and DeLeire, 2003; Heckman and Kautz, 2012; Mendolia and Walker, 2014b); health behaviors (e.g. Wallston *et al.*, 1978; Steptoe and Wardler, 2001; Chiteji, 2010; Cobb-Clark *et al.*, 2014; Mendolia and Walker, 2014a); employment-related training (Offerhaus, 2013; Caliendo *et al.*, 2016) and job search (e.g. Caliendo *et al.*, 2015b; McGee, 2015; McGee and McGee, 2016).[2] To our knowledge, however, only one other study explicitly considers the relationship between locus of control and migration. Toney *et al.* (1985) find no difference in the locus of control of migrant and non-migrant, middle-aged men captured in the U.S. National Longitudinal Survey. We reconsider this issue in a model that minimizes the likelihood of reverse causality and omitted variable bias. Our finding that having an internal locus of control is associated with a higher propensity to migrate across regions represents not only a new stylized fact, but also a potential basis for targeting internal migration incentives.

The outline of the paper is as follows. Section 3.2 provides an overview of the relevant economic and psychological literature on migration. In Section 3.3, the theoretical framework linking locus of control to job search and the migration decision is presented. Section 3.4 describes the data in detail, while in Section 3.5 we present our main empirical results. Section 3.6 provides a discussion of our main findings and Section 3.7 concludes the paper.

3.2 Literature Review

Economists have a long history of studying migration. Researchers taking a macro perspective typically analyze the relationship between migration flows and macro-economic conditions, while others adopt a micro perspective by focusing on single individuals (or households) and studying the migration decision-making process. Given our research questions, we are particularly interested in the micro-economics literature on internal migration and in the psychological evidence on the psychosocial traits that predispose certain individuals to migrate.

Drawing on the seminal work of Hicks (1932), Sjaastad (1962), Todaro (1969) and Harris and Todaro (1970), modern economic models of the migration decision are typically based on the maximization of expected income across regions, given migration costs and employment probabilities that are less than one. Seen in this light, it is easy to understand why economists view migration as an important form of human capital investment. Borjas *et al.* (1992) argue, however, that, while perhaps suitable for studying immigration, the above framework is too restrictive to capture internal migration because it predicts that migration flows will be unidirectional, i.e. all individuals have an incentive to move from low- to high-income regions. In response, the authors incorporate the Roy (1951) selection model into the migration decision

[2] For an overview of this literature see Cobb-Clark (2015).

thus accounting for spatial differences in the return to skill. This extension results in two-way migration flows with high-skilled workers migrating to regions where their skills are more highly valued and low-skilled workers migrating in the other direction. Conditional on skill level, however, migration flows remain unidirectional and, in the face of constant migration costs, this theoretical framework does little to explain why certain individuals might always be predisposed to migrate.

Given this framework, it is not surprising that empirical economists have largely focused on analyzing the way that migration patterns are shaped by skill levels, networks, migration costs and macro-economic conditions.[3] People's propensity to migrate has been linked to their age and gender (e.g. Stillwell *et al.*, 1996; Owen and Green, 1992); education and skill level (e.g. Levy and Wadycki, 1974; Arntz, 2010; Wozniak, 2010); individual social networks (e.g. Rainer and Siedler, 2009); marital status (e.g. Graves and Linneman, 1979); employment status (e.g. DaVanzo, 1978) and the business cycle (e.g. Saks and Wozniak, 2011). Internal migration is also a function of the costs of migration and regional disparities in social and economic circumstances as reflected in population size (density) and distance (e.g. Andrienko and Guriev, 2004; Anjomani, 2002; Greenwood, 1997), the cost of living (e.g. Cseres-Gergely, 2004), price differences (e.g. Giannetti, 2003); real wages (e.g. Kennan and Walker, 2011), unemployment rates (e.g. Ederveen *et al.*, 2003; Alecke *et al.*, 2010), labor demand (Wozniak, 2010), labor productivity (Alecke *et al.*, 2010), public safety (e.g. Sampson and Wooldredge, 1986), social assistance (e.g. Enchautegui, 1997; Giulietti and Wahba, 2013), climate and environmental quality (e.g. Marchiori and Schumacher, 2011; Andrienko and Guriev, 2004), and local infrastructure (e.g. Andrienko and Guriev, 2004).

Sociologists, demographers, and psychologists, on the other hand, all have a long tradition of studying the psychosocial traits that lead some individuals to be more likely to migrate irrespective of the economic conditions. Toney *et al.* (1985) attribute the first discussion of possible migrant-nonmigrant differences in psychological traits to Thomas (1938), a demographer and sociologist whose seminal work laid the foundation for migration research in the first half of the 1900s (see Greenwood and Hunt, 2003, for a discussion). Early researchers linked migration to a desire for social advancement (Touraine and Ragazzi, 1961) and the fulfillment of their achievement, affiliation, and power motivations (Frieze and Li, 2010). Morrison and Wheeler (1978) coined the term "pioneering personality" to describe individuals who constantly feel the need for novel experiences and thus like to change their residence. Since then researchers have found relationships between migration decisions and both economic preferences such as risk-attitudes (Jaeger *et al.*, 2010; Bauernschuster *et al.*, 2014; Bonin *et al.*, 2009) as well as personality traits such as openness to experiences (Koenig and Cunningham, 2001; Jokela, 2009), extraversion (Jacobs and Koeppel, 1974; Jokela, 2009) and agreeableness

[3] For reviews of the economics literature on internal migration see e.g. Borjas *et al.* (1992); Greenwood (1997); Lucas (1997); Etzo (2008).

(Jokela, 2009).[4] Finally, there is some evidence that an internal locus of control is associated with a modest increase in the willingness to move (Hines *et al.*, 1974), but not migration itself (Toney *et al.*, 1985).

In light of this empirical evidence, it is interesting that the migration literature is virtually silent on the role of job search *per se* in internal migration decisions. Early studies simply ignored any uncertainty associated with employment opportunities in either the sending or destination labor market rendering the migration decision a simple comparison of wage levels in the two locations. Since Todaro (1969) and Harris and Todaro (1970), it has become more common (though not universal) to assume that post-migration employment is not guaranteed. Kennan and Walker (2011), for example, model the optimal migration trajectory in the context of a dynamic search problem with multiple destination choices. Critically, however, migration is assumed to take place before destination wages are realized making the migration decision a function of expected income (see e.g. Treyz *et al.*, 1993; Borjas, 1999; Fuchs-Schündeln and Schündeln, 2009; Alecke *et al.*, 2010; Kennan and Walker, 2011). Effectively, the job search process boils down to a simple draw from the destination wage distribution.

In contrast, urban and labor economists view residential moves and job changes as being mutually dependent. Lutgen and Van der Linden (2015), for example, propose a model in which unemployed job seekers search across multiple regions and decide to migrate whenever they receive an acceptable job offer outside their local labor market. Models of job search thus incorporate the inherent tradeoffs associated with either commuting or moving in the event an acceptable job offer is received (see Rouwendal, 1999; van Ommeren *et al.*, 1999, 2000b; Eliasson *et al.*, 2003; Buchinsky *et al.*, 2014). Commuting time involves disutility, leading workers to trade off higher wages (den Berg and Gorter, 1997; van Ommeren *et al.*, 2000a) or make job changes (Zax, 1991; Zax and Kain, 1991) in exchange for lower commuting costs. Importantly, Zax (1994) shows that while job changes and residential moves can be substitutes in the case of intra-regional (local) mobility, they are most likely complements in the case of inter-regional (long-distance) mobility because commuting is not a viable option.

In what follows, we draw these strands of the literature together by incorporating locus of control into an economic model of internal migration that accounts for job search. The result is a more nuanced understanding of the process of internal migration, and the important role that job search and psychosocial traits like locus of control play in migration decisions.

3.3 Theoretical Framework

We begin with a conceptual framework in which households migrate from one geographic region to the next whenever the expected benefits of migration exceed the expected costs. We

[4] See Boneva and Frieze (2001); Frieze and Li (2010) for a review.

abstract from the choice of migration destination in order to focus on the discrete decision to stay or to go. Migration is modelled as a function of relative incomes, rather than relative utilities, in order to avoid the unnecessary complexity of considering migration based on regional amenities. Given our research focus, we pay particular attention to the benefits deriving from differences in labor market opportunities rather than from disparity in prices or social benefits.

Our interest is in understanding internal migration as the outcome of a job search process. However, we are agnostic about the relative timing of migration and job search, allowing for migration that occurs both before and after wages are realized. Geographic regions are assumed to be non-overlapping, ruling out commuting as a substitute for migration when inter-regional job changes occur. Finally, individuals are assumed to be rational. However, unlike standard job search models, we assume that individuals have subjective beliefs – related to their locus of control – about the impact of search effort on the job offer arrival rate.

This section proceeds as follows. Drawing on Borjas *et al.* (1992), we first discuss a standard model of internal migration which ignores job search.[5] We then extend this model to consider the implications of allowing potential migrants to engage in job search related to their locus of control. Finally, we consider the importance of the relative timing of job search and migration for the pattern of internal migration that results.

3.3.1 Internal Migration Ignoring Job Search

Following Borjas *et al.* (1992), individuals are assumed to have a single productive skill (x_i) which has return β_O in the origin labor market (O) and β_A in the alternative labor market (A). Thus, wages in the origin are given by: $w_{Oi}= \beta_O x_i$ and in the alternative region $w_{Ai} = \beta_A x_i$. Unlike Borjas *et al.* (1992), we assume that employment is uncertain and individuals receive a wage offer only with probability p_A in labor market A and p_O in labor market O. Households migrate whenever the net returns to migration are positive, that is whenever:

$$\begin{aligned} p_A w_{Ai} - p_O w_{Oi} - C &\geqslant 0 \\ (p_A \beta_A - p_O \beta_O) x_i - C &\geqslant 0, \end{aligned} \tag{3.1}$$

where C corresponds to a fixed cost of migration. Migration does not change individuals' skill levels, rather people migrate from O to A whenever the expected returns to their skill are higher in A (net of migration costs) than in O. Thus, the return to migration is generated by spatial differences in the returns to productive skills (β) and the probability of receiving a wage offer (p).

It is interesting to consider what this model implies about the nature of migration flows. As Borjas *et al.* (1992) note, migration flows are predicted to occur in two directions: highly-

[5] See also Hunt (2006); Fuchs-Schündeln and Schündeln (2009); Arntz *et al.* (2011) who adopt this framework.

skilled individuals have an incentive to migrate to regions in which skill is more highly valued and low-skilled individuals have an incentive to migrate in the opposite direction. Conditional on skill level (x_i) migration flows are unidirectional.[6] There are no barriers to internal migration as there would be across international borders. Thus, the internal migration of workers is expected to contribute to equalizing the return to skill across domestic labor markets until, in equilibrium, there is no incentive for further migration (see Borjas, 2000). Nothing in this simple framework explains why migrants with a particular skill x_i have a predisposition to migrate irrespective of their origin location. In the literature, bidirectional migration flows are usually introduced by idiosyncratic shocks. This often goes along with the assumption that these shocks are independent of observed characteristics and uncorrelated over time (see e.g. Nenov (2015) for a recent contribution that assumes idiosyncratic preference shocks for the migration decision).

3.3.2 A Model of Internal Migration with Spatial Job Search

We now consider a model in which employed and unemployed workers move to another region if they receive a job offer in that region which is above their reservation wage.

The probability of receiving a job offer depends on the job search effort s. Empirical researchers typically measure search intensity as: i) the number of applications made within a specific time frame (Caliendo *et al.*, 2015b); ii) the amount of time spent on search (McGee, 2015); or iii) the number of search channels utilized (Van Den Berg and Van Der Klaauw, 2006; Van Den Berg *et al.*, 2009). Instead, we argue that search effort can be characterized by the geographic distance D between the home region and the location of the potential employer.

This seems to us to be plausible. While the internet has reduced the information cost differential associated with searching in remote vs. local labor markets, workers will have additional knowledge about local firms and better access to local networks, job search agencies, and the like, all of which, everything else equal, facilitate local search. Searching over a greater geographic distance requires that workers increase their search effort. This implies that the geographic distance between the origin and alternative regions will be a function of search intensity with $\frac{\partial D}{\partial s} > 0$.[7]

Unlike in the standard job search model, we assume that individuals do not know the exact relationship between their own search effort s and the job offer arrival rate $\lambda(s)$. Instead, we assume that each person has a subjective belief about the impact of his or her search effort

[6] This is trivially true here because we consider only one type of skill, but would also be true if we allowed for a "skill profile". Conditional on each element of that skill profile, migration would be unidirectional.

[7] See Guglielminetti *et al.* (2015) who also develop a model of spatial job search in which search in remote areas is more costly. Job seekers take this into account when making decisions over their search intensity and search range. Based on recent data, Skandalis (2019) provides evidence that job seekers who apply for jobs in reaction to media stories that plants need to hire many workers live relatively far away from the growing plant, while job seekers living close to these plants do not apply for jobs in response to these stories with a higher probability. These findings suggest that information frictions increase with geographical distance.

on the job offer arrival rate. This subjective belief is characterized by individuals' locus of control, i.e. the degree to which they believe that there is a causal link between their own actions (search) and future outcomes (offer arrivals).

Subjective beliefs are given by $\tilde{\lambda}(s, loc)$, and those with an internal locus of control believe that an increased search effort results in a relatively large increase in the job offer arrival rate. The expected marginal return to search effort is therefore increasing in internal locus of control, i.e. $\frac{\partial^2 \tilde{\lambda}(s,loc)}{\partial s \partial loc} > 0$.

Caliendo *et al.* (2015b) present a job search model for unemployed workers which relies on the same assumption, while Ahn (2015) applies the same idea to an on-the-job search model for employed workers. Here we allow individuals to search for jobs both during unemployment and on the job. For simplicity, we assume that the search process does not differ between employed and unemployed workers. The job offer arrival rate $\lambda(s)$ and search costs $c(s)$ are the same for employed and unemployed workers. This implies that individuals accept every job offer with a wage above the unemployment benefit level b in case of unemployment and above the current wage w in case of employment. Therefore, the reservation wage simply corresponds to the current benefit level and wage, respectively.

Individuals choose their optimal search effort by equating the marginal costs of job search with the marginal benefits associated with additional search. The benefit of additional search is the increased probability of receiving a job offer paying more than the current wage or the unemployment benefit level, respectively. One can easily show that – because they expect a higher return to their search effort – individuals who have a more internal locus of control search more intensively than those with a more external locus of control (see Ahn, 2015; Caliendo *et al.*, 2015b):

$$\frac{\partial s^*}{\partial loc} > 0. \tag{3.2}$$

Given the relationship between search effort and geographic search area, this implies that individuals with a more internal locus of control will send more applications to other geographic regions than individuals with a more external locus control, i.e. $\frac{\partial D}{\partial loc} = \frac{\partial D}{\partial s^*} * \frac{\partial s^*}{\partial loc} > 0$. Although we cannot test this directly using SOEP data, we conduct an ancillary analysis of unemployed German job seekers using an alternative data source – the IZA Evaluation Dataset (Arni *et al.*, 2014).[8] We find in Table 3.B.2 that internal job seekers have an average maximum search distance (234 km) which is larger than that of their external counterparts (130 km). Moreover, estimation of a series of logit models with a rich set of controls reveals that internal job seekers are also significantly more likely to apply for jobs that would require a

[8] The IZA Evaluation Dataset contains survey information for a sample of individuals who entered unemployment in 2007 and 2008 in Germany. A nine percent random sample from the monthly unemployment inflows was selected for interview. In wave 1, 17,396 interviews were completed. A detailed discussion of our ancillary analysis including information about the IZA Evaluation Data, variable measurement, estimation strategy as well as descriptive statistics and full estimation results can be found in the Supplementry Appendix 3.B.

residential move (see Table 3.B.4). By searching more intensively across a broader geographic region, those with an internal locus of control are expected to be more likely to apply for – and be offered – jobs outside of their local labor markets.

3.3.3 Empirical Predictions

What are the empirical implications of incorporating locus of control into the internal migration decision through the wage returns that individuals expect? The answer depends critically on whether job search occurs before or after migration.

Migration before job search: Let us first assume that individuals migrate before destination wages are realized. Ex ante, the impact of any post-migration job search is simply captured in the aggregate employment probabilities conditional on their optimal search effort $p_A(s^*)$ and $p_O(s^*)$.[9] Because individuals' search intensity depends on their locus of control, these employment probabilities will also be a function of their locus of control. We can therefore rewrite equation (3.1) as follows:

$$\begin{aligned} p_A(s^*(loc))w_{Ai} - p_O(s^*(loc))w_{Oi} - C &\geqslant 0 \\ (p_A(s^*(loc))\beta_A - p_O(s^*(loc))\beta_O)x_i - C &\geqslant 0 \end{aligned} \tag{3.3}$$

In this simple case, workers migrate whenever – conditional on their skill level x_i – their expected wage gain from migrating exceeds the costs of doing so. In perfectly competitive labor markets, with perfect information and no mobility costs, the process of internal migration leads the returns to skill and employment probabilities to adjust so that in equilibrium expected wage differences across geographic locations equal migration costs. That is $p_A(s^*(loc))\beta_A - p_O(s^*(loc))\beta_O) = C$ and there is no further incentive for workers to migrate. Importantly, this process of labor market adjustment operates similarly in more realistic models that account for location-specific factors like the supply of land, local amenities, and housing costs in the migration decision. Roback (1982), for example, retains the assumption of perfect competition, but models internal migration in the context of identical workers with homogenous preferences over nominal wages, housing costs, and local amenities. Similarly, Moretti (2011) extends this framework by introducing idiosyncratic preferences for locations which also determine the degree of labor mobility. In both cases, internal migration is driven not by expected wages, but by a comparison of the indirect utility obtainable in different locations. As before, however, labor market competition ensures that the spatial movement

[9] Job search may also affect individuals' expected wage paths. Those who search more will have a faster wage growth over time, because they have a higher probability of finding a job with a higher wage than their current wage level. However, this will be the case in both labor markets A and O and so will not influence migration.

of workers continues until they are indifferent between locations, and there is no longer an incentive to migrate (Moretti, 2011).[10]

Thus, when migration occurs before job search and the realization of wages, the process of internal migration reduces the return to migration – whether measured in terms of expected wages or utility – until in equilibrium workers have no incentive to migrate. Consequently, there is nothing in this economic framework to explain multidirectional migration; that is, why some individuals might have a permanent incentive to migrate. Given this, our empirical prediction is that $P(M_{it}|\text{Internal}) = P(M_{it}|\text{External})$ where M_{it} is an indicator of migration for individual i at time point t. In contrast, many other social scientists believe that some individuals have personalities that predispose them to migrate (Touraine and Ragazzi, 1961; Morrison and Wheeler, 1978; Boneva and Frieze, 2001; Frieze and Li, 2010).

Migration after job search: Let us now consider internal migration that occurs only as the result of a successful job search. Because individuals with an internal locus of control perceive higher returns to their search effort, they search more intensively across a broader geographic area, i.e. $\frac{\partial D}{\partial loc} > 0$. The job offer arrival rate increases with search intensity implying that those with an internal locus of control will receive job offers in the alternative labor market more frequently than those with an external locus of control. Consequently, having an internal locus of control results in a greater propensity to migrate. When job search precedes migration, disparity in the search behavior of individuals with an internal locus of control predisposes them to migrate – even when the labor market is in equilibrium. Thus, we expect that $P(M_{it}|\text{Internal}) > P(M_{it}|\text{External})$.

Summary: Does migration occur before any job search or is the reverse true? The reality, of course, is that both can occur and we are unaware of any evidence indicating whether one dominates the other. In the case of immigrants and refugees, for example, both pull (e.g. family reunification) and push factors (e.g. natural disasters, persecution, civil wars, etc.) are almost certain to lead some individuals to migrate before they know their destination labor market outcomes. It is more likely, however, that internal migration is the consequence of a successful job search. van Ommeren *et al.* (1999), for example, note that while researchers typically make an assumption about the sequential ordering of the decision to change residences or to change jobs, there are sound theoretical reasons to expect that job moves trigger residential moves, but not the reverse. This is because there is a high probability that any increase in commuting costs associated with a job change can be reduced by a residential change. There is a much

[10] The theoretical prediction that internal migration contributes to eliminating regional disparities has empirical support. Blanchard *et al.* (1992), for example, find that the main mechanism that re-establishes labor market equilibrium after a demand shock appears to be labor mobility rather than job creation or job migration. Similarly, Topel (1986) carefully accounts for the dynamics of wage and employment changes and finds that positive shocks to labor demand increase nominal wages in a local labor market. Interestingly, this is not only true for current but also expected future demand shocks.

lower probability that any increase in commuting costs stemming from a residential move can be mitigated through a job change. We believe that this logic is particularly compelling in our case because we are explicitly focused on inter-regional migration involving substantial commuting costs across labor markets. Nonetheless, migration in our data will no doubt encompass both phenomena.

Therefore, given our theoretical framework, we expect there to be little relationship between migration and locus of control if migration predominately occurs before job search begins. We expect that an internal locus of control will be associated with an increased propensity to migrate across labor markets if job search prior to migration dominates. Consequently, we turn to the data to empirically discriminate between these alternatives.

3.4 Data

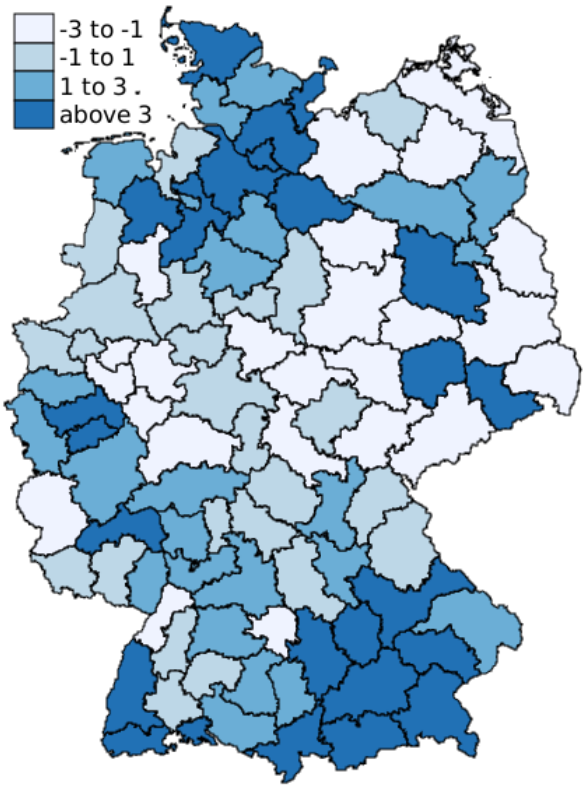
The data come from the German Socio-Economic Panel (SOEP) which is an annual representative panel study that collects detailed information about the socio-economic circumstances of approximately 22,000 individuals living in 12,000 households in Germany (see Wagner *et al.*, 2007, for details). These data are useful for our purposes because they provide measures of locus of control (and other personality traits) and identify the geographic location of the households in which individuals are living. Specifically, residential location is identified by geocodes which correspond to local planning regions ("Raumordnungsregionen") which broadly correspond to labor markets. This allows us to merge SOEP data with information on regional economic conditions, e.g. GDP, population density and unemployment rates.[11]

We restrict our study period to 1999-2015, as regional information are not available after 2015 and locus of control is first observed in 1999. Our population of interest is individuals between the ages of 25 and 55 who are not pensioners, on maternity leave, or in the military. We necessarily make a number of sample restrictions. Specifically, we exclude respondents with item non-response for the key variables of interest. We lose approximately six percent of our sample because we are unable to observe migration behavior, while item non-response in any one of the eight items underpinning the locus of control scale decreases our sample size by approximately one third.

We include all person-year observations available from the panel structure of the SOEP and consider this in our estimations using clustered standard errors. Thus, our estimation sample consists of 109,234 observations (53,141 men and 56,093 women) for 16,241 individuals (7,746 men and 8,495 women). See Table 3.A.1 for an overview of the sample loss associated with each selection criteria. Table 3.A.3 in the Appendix summarizes the key descriptive statistics for the sample.

[11] These data come from the INKAR database provided by the Federal Institute for Research on Building, Urban Affairs and Spatial Development.

Figure 3.1: Average Net Migration 1999-2015 in Local Labor Markets, per 1000 inhabitants



Source: INKAR 2018, own illustration

3.4.1 Measuring Internal Migration

Our indicator for internal migration is based on SOEP geocodes which allow us to classify each household's residential location into one of 96 separate regions. Although these regions do not correspond to official local government areas, they are the basis for the federal German government's regional planning. In particular, they capture urban centers (along with their associated catchment areas) and are defined on the basis of commuting flows (see BBSR, 2015). Researchers typically use the planning region as the unit of analysis when investigating issues such as geographic disparity in labor market conditions (e.g. Dütsch and Struck, 2014), employment growth (e.g. Fritsch and Noseleit, 2013) and regional mobility patterns (e.g. Jaeger *et al.*, 2010; Arntz, 2010). We use these regions to identify inter-regional mobility that corresponds to a change in labor markets. Specifically, our indicator of internal migration takes the value 1 if the household's geocode changes between t and $t+1$; and 0 otherwise.

Figure 3.1 depicts the average net migration flow (per 1,000 inhabitants) between regions in Germany over the period 1999-2015.[12] While most regions in East Germany are character-

[12] The pattern is similar if we rely on SOEP data rather than on the INKAR administrative data suggesting that our SOEP estimation sample is comparable.

ized by net out-migration (light shading), the major cities – such as Berlin, Munich, Hamburg and Frankfurt – as well as the surrounding areas are characterized by net in-migration (dark blue areas). This is consistent with previous findings. In particular, Arntz (2010) and Niebuhr *et al.* (2012) conclude that although the migration flow from East to West Germany began declining in 2001, it remains quite pronounced. Relative to other countries in the European Union, regional mobility in Germany is relatively high (Bonin *et al.*, 2008), though it is low in comparison to non-European countries such as the United States or Australia (Puhani, 2001; Bonin *et al.*, 2008).

In addition to actual migration, we also observe self-assessed willingness to move in 1999, 2009 and 2014. In these three years, respondents were asked *"Could you imagine moving away from here because of family or career reasons?"*. We use this information to create a binary indicator that takes the value 1 for individuals responding "yes"; and 0 for individuals responding "it depends" or "no". We observe self-assessed willingness to move for 11,033 (17,181) individuals (observations).

3.4.2 Measuring Locus of Control

In 1999, 2005, 2010 and 2015, SOEP respondents were asked how closely a series of 10 statements (items) characterized their views about the extent to which they influence what happens in life. A four-point Likert response scale ranging from 1 ('applies fully') to 4 ('does not apply') was used in 1999, while in 2005, 2010 and 2015 possible responses corresponded to seven-point Likert scale ranging from 1 ('disagree completely') to 7 ('agree completely'). A list of these items can be found in Table 3.A.2. In order to harmonize the scales, 1999 item responses are reversed and "stretched".[13] We conduct an explanatory factor analysis separately by year in order to investigate the way these items load onto latent factors.[14] The pattern of factor loadings is similar in all three years. Items 1 and 6 clearly load onto the first factor – which we interpret as internal locus of control – while items 2, 3, 5, 7, 8 and 10 clearly load onto the second factor – interpreted as external locus of control. Items 4 and 9 do not clearly load onto one factor or the other and are discarded.

Consistent with the previous literature (see, e.g., Piatek and Pinger, 2016; Cobb-Clark *et al.*, 2014), we use a two-step process to create a continuous locus of control index. First, we reverse the scores for items 2, 3, 5, 7, 8 and 10 so that all eight items are increasing in internal locus of control. Second, for each year, we use factor analysis to extract a single factor and mean standardize it. This has the advantage of allowing us to avoid simply weighting each item equally, as averaging would, and instead allow the data to drive how each item is weighted in the overall index. Simple averaging risks measurement error and attenuation

[13] This process preserves the relative differences between individuals. The process results in values of 1, 3, 5 or 7 so that a '1' on the 1999 four-point scale, for example, becomes a '7' on the 2005-2015 seven-point scales.
[14] The loading plots and the detailed results from the factor analysis are available upon request.

bias (Piatek and Pinger, 2016). Our locus-of-control index LOC_{it} is therefore increasing in internal locus of control and its distribution is shown in Figure 3.A.1 in the Appendix.[15]

There is evidence that locus of control is relatively stable for the working-age population (see e.g. Cobb-Clark, 2015; Cobb-Clark and Schurer, 2013). Nevertheless, in order to minimize concerns about potential reverse causality, we ensure that our locus of control index is always measured prior to the period in which we observe the migration decision. That is, migration in 2000-2004 depends on 1999 locus of control, migration in 2006-2009 depends on 2005 locus of control, migration in 2011-2014 depends on 2010 locus of control and migration 2015-2016 depends on 2015 locus of control. In addition to the continuous measure, we also create an indicator of "internal" locus of control which takes the value 1 for those with locus of control indexes above the median; and 0 otherwise. Finally, we test the robustness of our results to different specifications of this indicator in Section 3.5.3.

3.4.3 Locus of Control and Internal Migration

Overall, 1.5 percent of the individuals in our sample moved across regions between t and $t+1$ with men (1.5 percent) being slightly more likely to migrate than women (1.4 percent)(see Table 3.1). Moreover, one in four individuals (25.1 percent) report that they are definitely willing to migrate for family or career reasons, while a further 43.0 percent report that they would consider migrating under some circumstances.

Those with an internal locus of control are significantly more likely (1.7 percent) to move across regions than are those who are external (1.3 percent). This pattern is also supported by the kernel densities computed separately for movers and non-movers in Figure 3.A.1 (bottom panel). Similarly, those with an internal locus of control are significantly more likely to express a willingness to migrate and significantly less likely to rule migration out.

Is an expressed willingness to migrate related to actual migration behavior? We consider this question and find that both men and women who do in fact migrate between t and $t+1$ are significantly more likely to have reported a willingness to migrate in t. Specifically, while only 24.0 percent of non-migrants report a willingness to migrate, nearly 71.0 percent of migrants did the same in the period prior to their move (see Table 3.A.3 in the Appendix).

3.5 Empirical Approach and Results

3.5.1 Estimation Strategy

Our theoretical model predicts that internal migration and internal locus of control will be positively related if job search precedes migration and unrelated if job search follows migration. We have no prior about the relative timing of job search and migration, making

[15] A test of internal consistency yields a Cronbach's α reliability statistic (Cronbach, 1951) between 0.66 and 0.68 indicating that the eight items are reliable which is in line with previous studies (Richter *et al.*, 2013).

Table 3.1: Descriptives of Outcome Variables by Locus of Control

	All	Externals	Internals
All			
Internal Migration between t and $t+1$	0.0146	0.0127	0.0165***
	[109,234]	[54,720]	[54,514]
Could Imagine Moving Away			
No	0.3194	0.3291	0.3098***
It Depends	0.4301	0.4370	0.4232*
Yes	0.2505	0.2339	0.2670***
	[17,181]	[8,559]	[8,622]
Men			
Internal Migration between t and $t+1$	0.0148	0.0129	0.0166***
	[53,141]	[25,599]	[27,542]
Could Imagine Moving Away			
No	0.3121	0.3229	0.3022***
It Depends	0.4272	0.4372	0.4179*
Yes	0.2607	0.2399	0.2799**
	[8,362]	[4,014]	[4,348]
Women			
Internal Migration between t and $t+1$	0.0144	0.0125	0.0165 ***
	[56,093]	[29,121]	[26,972]
Could Imagine Moving Away			
No	0.3263	0.3347	0.3175 ***
It Depends	0.4328	0.4367	0.4286
Yes	0.2408	0.2286	0.2539 *
	[8,819]	[4,545]	[4,274]

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Number of observations in square brackets. Significance stars in the last column refer to significance level of t-test for mean equivalence between Externals and Internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the link between locus of control and migration an empirical question. Consequently, we employ a reduced-form approach to estimate the association between individuals' locus of control and their propensity to: i) express a willingness to migrate; and ii) to actually migrate across regions. Specifically, our estimation equations are as follows:

$$P(W_{it} = 1) = P(\theta_1 + \theta_2 LOC_{it} + \theta_3 X_{it} + \theta_5 PT_i + \tau_1 R_{it} + \tau_2 T + \epsilon_{it} > 0) \quad (3.4)$$

$$P(M_{it+1} = 1) = P(\beta_1 + \beta_2 LOC_{it} + \beta_3 X_{it} + \beta_5 PT_i + \gamma_1 R_{it} + \gamma_2 T + \eta_{it} > 0) \quad (3.5)$$

where i indexes individuals, t indexes time, and W_{it} and M_{it+1} capture the stated willingness to migrate (at time t) and actual migration behavior (between t and $t+1$) respectively. Further, LOC_{it} is locus of control and θ_2 and β_2 are our parameters of interest. We generate interpretable estimates of θ_2 and β_2 by constructing our locus of control measure such that it is predetermined at the time of the migration decision in order to minimize concerns about

reverse causality. In addition, we include a detailed set of controls to reduce the potential for unobserved heterogeneity (omitted variable bias) to confound our estimates. Hence, X_{it} includes standard socio-demographic characteristics (such as gender, age, nationality, marital status, number of children, household income, home ownership, and disability status) as well as controls for education (school degree, vocational education, university degree) and job characteristics (current labor force status, occupational classification, tenure and unemployment experience). PT_{it} is a vector of individual personality traits averaged over all years (Big Five traits and risk attitudes). Finally, R_{it} captures regional conditions (dummy for East Germany, unemployment rates, gross value added and population density in the origin region) and T is a vector of year-dummies.

All these factors have been shown to be important in explaining internal migration (see e.g. Kennan and Walker, 2011; Ederveen *et al.*, 2003; Jokela, 2009; Alecke *et al.*, 2010; Jaeger *et al.*, 2010). Some of our controls (e.g. employment history or education) may themselves be a function of locus of control. We will investigate the robustness of our results by leaving out these potentially endogenous regressors from our regressions in Section 3.5.3. Descriptive statistics for the variables in our analysis are reported in Table 3.A.3 in the Appendix.

Equations 3.4 and 3.5 are estimated using logit models with standard errors clustered at the person level. All estimated effects are presented as average marginal effects in percentage points. We estimate three alternative specifications. The first controls only for year and regional indicators (T and R_{it}). The second adds controls for socioeconomic and job characteristics (X_{it}), while the third also controls for Big-Five personality traits and risk attitudes (PT_{it}). Models are estimated separately by gender using one of two alternative measures of locus of control: i) a continuous measure; and ii) an indicator for having a locus of control greater than the median, i.e. being internal. Table 3.2 (willingness to move) and Table 3.3 (actual internal migration) provide an overview of the key results, while full estimation results for our preferred specification – including all groups of control variables – are available in Tables 3.A.4 and 3.A.5 in the Appendix.

3.5.2 Locus of Control and Internal Migration

Individuals with an internal locus of control are more likely to report that they would consider moving for family or career reasons. Moreover, this relationship is robust to the inclusion of a detailed set of controls (see Table 3.2). Controlling for personality traits and risk aversion leads to an increase in the estimated effects of locus of control (see columns (3) vs. (5) and (4) vs. (6) respectively).

Specifically, each standard deviation increase in individuals' internal control tendencies results in a one percentage point (p.p.) increase in the likelihood that individuals respond "yes" when asked if they are willing to migrate (see column 5). Those with an internal locus of control, i.e. those above the median, are 2.5 percentage points more likely to report

Table 3.2: Logit Estimation Results (Marginal Effects): Willingness to Move

	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
All Individuals						
LOC Factor (std.)	1.171***		0.859**		1.030***	
	(0.351)		(0.354)		(0.369)	
LOC Factor > Median		2.910***		2.052***		2.462***
		(0.685)		(0.684)		(0.695)
Observations	17,181	17,181	17,181	17,181	17,181	17,181
Pseudo R^2	0.015	0.015	0.057	0.057	0.065	0.065
Men						
LOC Factor (std.)	1.742***		1.380***		1.678***	
	(0.508)		(0.515)		(0.543)	
LOC Factor > Median		3.759***		2.885***		3.443***
		(0.998)		(0.993)		(1.008)
Observations	8,362	8,362	8,362	8,362	8,362	8,362
Pseudo R^2	0.024	0.024	0.071	0.071	0.081	0.081
Women						
LOC Factor (std.)	0.545		0.387		0.450	
	(0.484)		(0.485)		(0.501)	
LOC Factor > Median		1.941**		1.320		1.629*
		(0.938)		(0.937)		(0.958)
Observations	8,819	8,819	8,819	8,819	8,819	8,819
Pseudo R^2	0.008	0.008	0.052	0.052	0.059	0.060
Year Fixed-Effects	✓	✓	✓	✓	✓	✓
Regional Controls	✓	✓	✓	✓	✓	✓
Socio-Economic Controls			✓	✓	✓	✓
Personality Controls					✓	✓

Source: Socio-Economic Panel (SOEP) 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level.
Full estimation results for specification (6) can be found in table 3.A.4, in Appendix 3.A. The full estimation results for specifications (1) - (5) can be obtained from the authors.

being willing to move relative to those with an external locus of control (see column 6). Unfortunately, previous evidence linking locus of control to a willingness to migrate is virtually nonexistent making it difficult to compare results. The exception is early research by Hines *et al.* (1974) who also find a positive correlation between internal locus of control and self-assessed willingness to migrate in a very small sample (n=53) of undergraduate students. Our results provide evidence that this finding is pervasive in a much broader population. Overall, 25.1 percent of our estimation sample reports being prepared to migrate implying that the disparity associated with locus control amounts to a difference of approximately 10.0 percent. This is of the same order of magnitude as having an university degree (3.3 p.p), being a white-collar worker (2.3 p.p.), or being married (-3.1 p.p) and is larger than the effect associated with having an additional 1,000 Euro in household income (0.4 p.p) (see Table 3.A.4 in the Appendix for full estimation results). The relationship between willingness to move and locus of control is generally stronger for men than for women. Men with an internal locus of control, i.e. those above the median, are 3.4 percentage points more likely to report being willing to move (relative to those with an external locus of control) while the effect for women is only 1.6 percentage points.

Turning to Table 3.3, we see that individuals with an internal locus of control are also more likely to migrate between labor market regions from one year to the next than are those with an external locus of control. Specifically, individuals who are internal have a 0.2 percentage point higher probability of moving each year than do external individuals (column 6). While small, this effect is economically meaningful given that the annual rate of internal migration on average is only 1.5 percent. Thus, the estimated effect of an internal locus of control translates into a 13.0 percent higher probability of moving.

This is comparable to the effect of an additional child in the household (-0.3 p.p.) and is larger than the effect of two more years of unemployment experience (-0.2 p.p.) (see Table 3.A.5). Our continuous measure of locus of control is also positively associated with increased migration (column 5), however, this association is not quite significant at conventional levels. These findings are in contrast to those of Toney *et al.* (1985) who find no relationship between geographical mobility and locus of control for middle-aged, white men captured in the U.S. National Longitudinal Survey.

There are several important things to note about these empirical results. First, the relationship between internal migration and locus of control is largely unaffected by the inclusion of a wide range of additional controls. For example, controlling for personality traits (Big Five) and risk aversion does not change the estimated effect of the binary indicator for the full sample (column 6 vs. column 4). The effect for the continuous measure slightly decreases and becomes insignificant, but is qualitatively very similar (column 5 vs. column 3). For the sample of men the point estimates are even increasing after controlling for personality traits and risk aversion, while they are decreasing in the sample of women. Thus, locus of control has

Table 3.3: Logit Estimation Results (Marginal Effects): Internal Migration between t and $t+1$

	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
All Individuals						
LOC Factor (std.)	0.152***		0.072*		0.064	
	(0.042)		(0.043)		(0.046)	
LOC Factor > Median		0.365***		0.189**		0.189**
		(0.082)		(0.081)		(0.083)
Observations	109,234	109,234	109,234	109,234	109,234	109,234
Pseudo R^2	0.004	0.004	0.127	0.127	0.130	0.130
Men						
LOC Factor (std.)	0.123**		0.067		0.077	
	(0.062)		(0.066)		(0.070)	
LOC Factor > Median		0.355***		0.225*		0.250**
		(0.120)		(0.123)		(0.126)
Observations	53,141	53,141	53,141	53,141	53,141	53,141
Pseudo R^2	0.005	0.005	0.118	0.118	0.120	0.121
Women						
LOC Factor (std.)	0.172***		0.075		0.052	
	(0.056)		(0.056)		(0.061)	
LOC Factor > Median		0.365***		0.150		0.125
		(0.113)		(0.108)		(0.111)
Observations	56,093	56,093	56,093	56,093	56,093	56,093
Pseudo R^2	0.006	0.006	0.147	0.147	0.150	0.151
Year Fixed-Effects	✓	✓	✓	✓	✓	✓
Regional Controls	✓	✓	✓	✓	✓	✓
Socio-Economic Controls			✓	✓	✓	✓
Personality Controls					✓	✓

Source: Socio-Economic Panel (SOEP) 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level.
Full estimation results for specification (6) can be found in table 3.A.5, in Appendix 3.A. The full estimation results for specifications (1) - (5) can be obtained from the authors.

important additional explanatory power in models of the migration decision over and above that associated with both personality traits and more traditional economic drivers including job characteristics, regional economic conditions, family structure, and preference parameters (e.g. risk-attitudes).

Secondly, the overall relationship between internal migration and locus of control is largely driven by men. Men with an internal locus of control are not only more likely to report a willingness to move than are men with an external locus of control, they are on average also 17.3 percent (0.3 p.p.) more likely to migrate than external men. At the same time locus of control is not significantly related to the migration behavior of women.[16][17] These gender differences are consistent with a growing literature showing that there is a gender-specific relationship between locus of control and many labor market outcomes including wages (Semykina and Linz, 2007), occupational attainment (Cobb-Clark and Tan, 2011), job search (Caliendo *et al.*, 2015b), employment-related training (Offerhaus, 2013; Caliendo *et al.*, 2016), selection into jobs with performance appraisal (Heywood *et al.*, 2017), and entrepreneurship (Hansemark, 2003).

Our results are consistent with job search preceding internal migration for the majority of individuals. In particular, our theoretical model predicts that locus of control will be unrelated to internal migration if individuals migrate prior to engaging in job search. In contrast, we find a positive relationship between the propensity to migrate and having an internal locus of control suggesting that migration for many individuals results from successful job search. Interestingly, van Ommeren *et al.* (1999) reach a similar conclusion arguing that it is theoretically more plausible that job search precedes migration than the reverse.

The relationships between migration and the other independent variables are very much in line with prior expectations and the earlier literature (see Tables 3.A.4 and 3.A.5). Age has a significant U-shaped effect on willingness to move and a significantly negative effect on the probability of internal migration. Although the effect of locus of control on internal migration differs by gender, men are not, ceteris paribus, significantly more likely to migrate. Rather being married, being a home-owner or having more than three years of job tenure are all associated with significantly less willingness to move and lower propensities to actually migrate. On the other hand, self-assessed willingness to move as well as actual internal migration are significantly higher for the more educated, the unemployed, white-collar workers as well as people in West Germany. Interestingly, the probability of internal migration is

[16] Interestingly, in our anciliary analysis of unemployed job-seekers where the effect of locus of control is stronger on the migration decisions of women than those of men (see Table 3.B.4). This is likely driven by the selectivity of women into the IZA sample. Whereas 66 percent of the women in the SOEP are married/live with a partner, this is only true for 44 percent of women in the IZA data, making them less likely to be "tied movers".

[17] However, in a fully interacted model the coefficient on $LOC \times Gender$ is not significant (neither in the SOEP nor the IZA sample). Results available on request from the authors.

negatively related to regional gross value added, while individuals' self-assessed willingness to move increases with regional population density.

3.5.3 Sensitivity Analysis

We conducted a sensitivity analysis in order to test the robustness of our results to: i) different measures of locus of control, ii) controls for willingness to move, iii) the exclusion of potentially endogenous variables and iv) controls for earlier moves. Our focus is on the estimated association of our dichotomous measure of internal locus of control with actual migration decisions. All sensitivity results are summarized in Table 3.4 and are based on the full specification (column 6 in Table 3.3).

We begin by considering whether our estimates are driven by the choice of the median as the threshold for identifying those with an internal locus of control. Specifically, we re-estimated our models using: i) mean locus of control as a threshold; and ii) a four-way classification of internal locus of control, namely low ($< 25^{th}$ percentile), lower medium (25^{th} - 50^{th} percentile), upper medium (50^{th} - 75^{th} percentile) and high ($> 75^{th}$ percentile). Our results are robust to these alternative measures (see Panel A, (1) and (2) in Table 3.4). Moreover, we replaced our preferred measure of locus of control which relies on factor weights with an alternative based on an equal weighting (simple average) of the underlying eight locus of control items. We find that our results continue to hold using this alternative measure (Panel A, (3) in Table 3.4). Finally, although it can be argued that locus of control is relatively stable in adulthood, we nevertheless want to investigate whether our results hold if we rely only on between individual rather then also within individual variation in locus of control. Consequently, we re-estimate our models using the first available measure of locus of control for each individual, i.e. 1999 or later if individuals enter the sample after 1999 or have a missing locus of control in 1999. The estimated effects using this alternative imputation are even more pronounced than in the baseline specification (see Panel A (4)).

Next we consider the following question: To what extent is the relationship between internal locus of control and internal migration operating through a heightened willingness to move? We consider this question by including a control for self-assessed willingness to move in our model of actual migration behavior (see Panel B in Table 3.4). Unfortunately, willingness to move is only observed in 1999, 2009 and 2014. Therefore, to retain as much sample as possible we impute self-assessed willingness to move between observation periods using the most-recently available measure. This effectively requires us to maintain the strong assumption that willingness to move is stable across years. We find that, not surprisingly, willingness to migrate is closely linked to actual migration behavior. Individuals who report a willingness to move have a probability of actually migrating that is fully 1.1 percentage points higher. Nonetheless, controlling for willingness to move has only a marginal effect on the magnitude of the association between our indicator of internal locus of control and internal migration.

Table 3.4: Robustness Checks (Marginal Effects): Internal Migration between t and $t+1$

	All	Men	Women
Baseline Results			
LOC Factor > Median	0.189**	0.251**	0.126
	(0.083)	(0.127)	(0.111)
Observations	109,234	53,141	56,093
A. Alternative LOC Specifications			
(1) Alternative Dichotomous Variable			
LOC Factor > Mean	0.177**	0.268**	0.124
	(0.083)	(0.125)	(0.111)
Observations	109,234	53,141	56,093
(2) Finer Distinction			
$(LOC_{P25}, LOC_{P50}]$	0.052	-0.014	0.136
	(0.111)	(0.161)	(0.154)
$(LOC_{P50}, LOC_{P75}]$	0.226*	0.238	0.152
	(0.118)	(0.179)	(0.153)
$(LOC_{P75}, LOC_{max}]$	0.209*	0.247	0.262
	(0.118)	(0.180)	(0.159)
Observations	109,234	53,141	56,093
(3) Simple Index LOC Calculation[a]			
LOC Index > Median	0.170**	0.256**	0.082
	(0.085)	(0.126)	(0.113)
Observations	109,234	53,141	56,093
(4) First LOC Observation[b]			
LOC Index > Median	0.212**	0.365***	0.111
	(0.085)	(0.125)	(0.116)
Observations	109,234	53,141	56,093
B. Willingness to Move as Intermediate[c]			
LOC Factor > Median	0.153*	0.241*	0.071
	(0.084)	(0.127)	(0.113)
Dummy for Willingness to Move (imp.)	1.130***	0.925***	1.287***
	(0.093)	(0.133)	(0.131)
Observations	105,446	51,027	54,222
C. Excluding Potentially Endogenous Control Variables[d]			
LOC Factor > Median	0.303***	0.389***	0.212*
	(0.082)	(0.125)	(0.112)
Observations	109,234	51,224	54,222
D. Control for Earlier Moves			
LOC Factor > Median	0.164**	0.243**	0.086
	(0.080)	(0.121)	(0.106)
Number of Earlier Moves	0.605***	0.665***	0.523***
	(0.043)	(0.066)	(0.057)
Observations	109,234	53,141	56,093
Year Fixed-Effects	✓	✓	✓
Regional Controls	✓	✓	✓
Socio-Economic Controls	✓	✓	✓
Personality Controls	✓	✓	✓

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level. All rows of marginal effects and standard errors are from separate estimations.
[a] The Simple Index is calculated by averaging over the item values in the following way: [I1 + I6 + R(I2 + I3 + I5 + I7 + I8 + I10)]/8 where R indicates that all external items are reversely coded.
[b] The LOC is imputed forward from the first LOC observation available for the individual in the SOEP, i.e. 1999 or later if individuals enter the sample after 1999 or have a missing LOC in 1999.
[c] The Willingness to Move variable is imputed forward in order to have a valid observation for all years in the full sample. The observation from 1999 is imputed into the years 2000 - 2008, the observation from 2009 is imputed into the years 2010 - 2013 and the observation from 2014 is imputed into the year 2015.
[d] In this specification we exclude all controls for education and labor-market situation and history.

We find similar results in a specification that uses a three-way categorical measure of the willingness to move, rather than a simple indicator.

Additionally, a large literature demonstrates that personality traits, including locus of control, are related to individuals' human capital acquisition (see, e.g., Coleman and DeLeire, 2003; Heckman *et al.*, 2006). Thus, many of our human capital measures, in particular education and employment histories, may themselves be a function of individuals' perceptions of control via decisions made in the past. The inclusion of these endogenous controls in the analysis is likely to introduce selectivity bias (see e.g. Angrist and Pischke, 2008, p.34). To gauge the importance of these endogenous controls, we re-estimate our models excluding educational attainment and employment information from the regression (Panel C in Table 3.4). We find that the relationship between internal migration and internal locus of control becomes stronger. Specifically, the estimated association of internal locus of control with actual migration behavior increases from 0.2 to 0.3 percentage points, which is equivalent to an increase from 13 percent to 20.8 percent. For men the increase is from 0.3 to 0.4 percentage points (17.2 percent to 26.3 percent) and for women the increase is from 0.1 to 0.2 percentage points (from 8.8 percent to 14.7 percent), making the association significant.

Finally, we investigate whether our results change when we control for previous migration. Having moved once in life probably makes a second move more likely. This is why we include the number of earlier moves as a control into the model (see Panel D in Table 3.4).[18] As expected, having moved before is associated with a positive increase in the probability of moving again. However, the estimated effect for LOC is stable and only changes slightly. Previous migration behavior is likely to be correlated with unobserved characteristics which affect the migration decision and which are potentially correlated with locus of control. Therefore, the stability of our results makes us confident that locus of control is not only reflecting differences in other unobserved characteristics. However, previous migration behavior is obviously an endogenous variable, and therefore we prefer to not include these in our main specifications.

Taken together, these sensitivity tests indicate that our estimates of the positive relationship between internal migration and internal locus of control are robust to a range of specification issues.

3.6 Discussion

There is substantial evidence that perceptions of control are related to human capital investment decisions through the returns that individuals expect (see Cobb-Clark, 2015). This makes it quite natural to link locus of control to internal migration through the expected

[18] Earlier moves refer to moves between regions while being observed in the SOEP. Consequently, the variable contains the total number of observed moves between regions for an individual while being observed in the SOEP.

Table 3.5: Descriptives: Context of Moving

	(1) All	(2) Internals	(3) Externals
(1) Difficulty Level of Finding an Appropriate Position			
Easy	0.19	0.24	0.15***
Difficult	0.61	0.59	0.62***
Impossible	0.16	0.14	0.19***
Observations	109,234	54,514	54,720
(2) Reason of Moving - Sample: All Movers[a]			
Work Reasons	0.13	0.15	0.12***
Observations[b]	5,611	2,823	2,788
(3a) Moving Distance - Sample: All Movers[a]			
Distance in km	31.71	36.16	27.13***
Observations	5,893	2,987	2,906
(3b) Share of Regional Movers - Sample: All Movers[a]			
Inter-regional Move	0.17	0.19	0.15***
Observations	5,893	2,987	2,906

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: All numbers are shares unless stated otherwise. Significance stars in the last column refer to significance level of t-test for mean equivalence between Externals and Internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[a] The sample consists of all individuals who indicate that they moved residence in the interview in $t+1$, including all inter- and intraregional moves.

returns to job search as we have done here. This conceptual link is also consistent with descriptive evidence that those with an internal locus of control are significantly: i) more likely to view finding a new job as "easy"; ii) less likely to believe finding a new job will be "difficult" or "impossible";[19] and iii) more likely to migrate for work-related reasons[20] (see Table 3.5). Migrants with an internal locus of control are significantly more likely to make inter-regional moves, leaving them to migrate over significantly greater distances on average.[21]

Despite the logic of the endeavor, however, we cannot completely rule out the possibility that locus of control may operate through some other mechanism. Our theoretical predictions are observationally equivalent, for example, if we instead assume that having an internal locus of control raises the efficiency of job search; if job search precedes migration, those with an internal locus of control will have a higher propensity to migrate.

At the same time, our results are not consistent with a model in which locus of control operates solely through labor market productivity, i.e. through wages. Were this the case, internal migration would be driven by differences in the return to control perceptions across

[19] The question is asked for employed/self-employed as well as unemployed individuals in a nearly identical way. While unemployed individuals are asked "If you were currently looking for a new job: Is it or would it be easy, difficult, or almost impossible to find an appropriate position?", all other individuals respond to the question "If you lost your job today, would it be easy, difficult, or almost impossible for you to find a new job that is at least as good as your current one?"

[20] Individuals who moved residence are asked to indicate what the reason for their move was with "work reasons" being one of the multiple options.

[21] Moving Distance is a generated variable provided by the SOEP which is based on the address data of the new and old residential location.

labor markets. Those with an internal locus of control would be pulled towards markets with high returns to being internal, while those with an external locus of control would be pulled in the opposite direction. This internal migration would contribute to equalizing the return to locus of control across labor markets until, in equilibrium, we would expect no relationship between locus of control and migration.

Nor does it seem likely that locus of control simply captures the effects of lower migration costs. There is very little reason to believe that the monetary costs of migration (e.g. moving expenses, buying and selling houses, etc.) are related to individuals' locus of control. Bauernschuster *et al.* (2014), however, argue that lower psychic migration costs lie behind the higher migration rates of highly-educated and risk-tolerant individuals, while Moretti (2011) models worker heterogeneity in preferences for location which can be conceptualized in our context as a psychic (opportunity) cost of migration. If internal locus of control affects internal migration by lowering the psychic costs of migration, internal individuals will have a greater incentive to migrate regardless of the relative timing of migration and job search. However, to the extent that self-assessed willingness to move captures heterogeneity in the psychic costs of migration, our empirical results imply that differences in psychic migration costs do not provide a complete explanation for predisposition of those with an internal locus of control to migrate across labor markets.

Future research which assessed the empirical support for these alternative explanations and the apparent differences in the behavior of men and women would be useful in furthering our understanding of the mechanisms linking locus of control and internal migration.

3.7 Conclusions

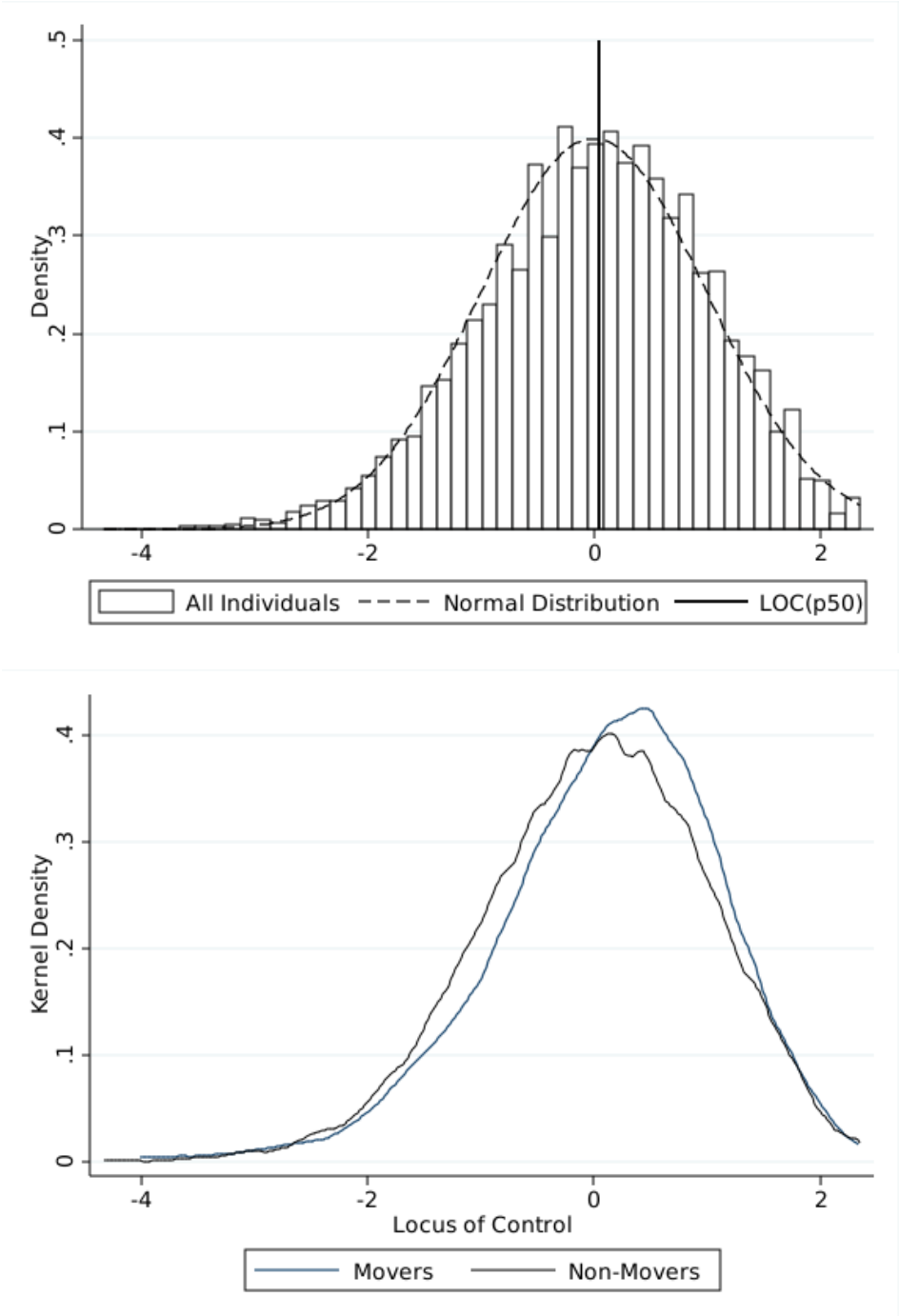
Internal migration is intrinsically linked to economic opportunity, social conditions, and demographic change, making it one of the most commonly studied phenomenon in the social sciences. Our objective is to develop a more nuanced understanding of the process of internal migration by drawing together several strands of the literature. Specifically, we model internal migration as the result of a job search process and demonstrate that the relative timing of job search, i.e. either before or after migration, has important implications for the nature of migration flows. When migration precedes job search, migration flows will be unidirectional conditional on skill as in Borjas *et al.* (1992). High-skilled workers will have an incentive to move to markets with high returns to skills, while low-skilled workers tend to migrate in the opposite direction. On the other hand, when job search precedes migration, those characteristics, skills, traits, etc. that facilitate geographically broader job search will be associated with a higher propensity to migrate.

This relationship between internal migration and job search provides an important conceptual framework for understanding the psychological evidence that certain traits predispose individuals to migrate. In particular, we explicitly model individuals' subjective beliefs about

the returns to job search as a function of their locus of control. Those with an internal locus of control search more intensively, across a wider geographic area, and therefore have a higher propensity to migrate. We then test the empirical implications of our model and find that those with an internal locus of control both report being more willing to migrate and in fact do migrate more often. This positive relationship between locus of control and internal migration constitutes a new stylized fact and indicates that, for many individuals, job search is likely to precede migration. Given this, providing incentives for more intensive job search across a wider geographical area – particularly when targeted towards those with an external locus of control – may be more useful in increasing internal migration rates than are standard relocation initiatives.

3.A Appendix

Figure 3.A.1: Distribution of Locus of Control



Source: SOEP waves 1999-2015, version 33, own illustration.
Notes: Bottom panel provides the graphical illustration of the kernel density estimates using the Epanechnikov kernel function separately for movers and non-movers.

Table 3.A.1: Sample Selection and Item Non-Response

Step	Estimation Sample	
	Observations	Individuals
Full Sample (1999-2015)	647,304	100,409
Sample Restriction		
- Drop Younger 25, Older 55	279,956	48,833
- Drop Pensioneers, Mat. Leave, Milit. Service	268,190	48,347
Item Non-Response		
- Migration Variable	244,087	45,570
- Locus of Control	138,100	27,419
- Socio-Economic Controls	122,562	25,498
- Personality Control	109,234	16,241
Sub-Sample for Willingness to Move (1999, 2009 & 2014)[a]	17,181	11,003

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: The full sample contains all available SOEP observations between 1999 and 2015 including e.g. children and persons without person questionnaires.
[a] The number of observations includes all sample restrictions and item non-responses.

Table 3.A.2: Components of Locus of Control

	Wave	
Variable	1999[a]	05/10/15[b]
Components of locus of control (Mean, 1999 Scale: 1-4, 2005/10/15 Scale: 1-7)		
I1: How my life goes depends on me	3.28	5.50
I2: Compared to other people, I have not achieved what I deserve (R)	2.12	3.25
I3: What a person achieves in life is above all a question of fate or luck (R)	2.21	3.42
I4: If one is soc. or polit. active, one can have an effect on social conditions[c]	2.29	3.69
I5: I freq. have the experience that others have a controlling influence over my life (R)	2.00	3.14
I6: One has to work hard in order to succeed	3.46	5.93
I7: If I run up against difficulties in life, I often doubt my own abilities (R)	2.03	3.26
I8: Opportunities I have in life are determined by the social conditions (R)	2.69	4.48
I9: Inborn abilities are more important than any efforts one can make[c]	2.93	4.78
I10: I have little control over the things that happen in my life (R)	1.80	2.61
Observations	5,419	23,276

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Items marked with a (R) are reversed prior to factor analysis in order to indicate an internal locus of control for high values.
[a] In 1999 the LOC was surveyed on a 4-point Likert scale from 1 for "Totally Disagree" to 4 for 'Totally Agree". The scale was reversed in the data preparation in order to indicate agreement for high values as it is also the case in the other waves. For the later harmonization, the scale was stretched to the length of a 7-point Likert scale.
[b] In 2005 and 2010 the LOC was surveyed on a 7-point Likert scale from 1 for "I Disagree Completely" to 7 for "Agree completely".
[c] Items 4 and 9 are not included into the prediction of the latent factor.

Table 3.A.3: Selected Descriptive Statistics by Internal Migration Status

	All		Male		Female	
	Non-Migrants	Migrants	Non-Migrants	Migrants	Non-Migrants	Migrants
Willingness to Migrate[a]	0.24	0.71***	0.25	0.66***	0.23	0.75 ***
Locus of Control[b]	4.33	4.44***	4.38	4.46**	4.29	4.41 ***
Socio-Economic Characteristics						
Age	41.39	34.55***	41.29	34.74***	41.48	34.36 ***
Married	0.66	0.33***	0.65	0.36***	0.67	0.30 ***
Number of Children in HH	0.73	0.42***	0.73	0.42***	0.72	0.43 ***
German	0.92	0.95***	0.92	0.94**	0.92	0.95 ***
East-Germany	0.27	0.24**	0.27	0.23***	0.27	0.26
Disabled	0.06	0.04***	0.07	0.04***	0.05	0.03 **
Home-Owner	0.52	0.21***	0.52	0.23***	0.51	0.19 ***
School Degree						
No Degree	0.02	0.01**	0.02	0.01	0.02	0.01 *
Lower Secondary School	0.24	0.14***	0.28	0.17***	0.21	0.11 ***
Intermediate School	0.37	0.28***	0.33	0.27***	0.41	0.29 ***
Highschool	0.30	0.55***	0.32	0.52***	0.29	0.57 ***
Vocational Education						
No Vocational Education	0.28	0.42***	0.27	0.40***	0.29	0.43 ***
Apprenticeship	0.45	0.36***	0.48	0.41***	0.43	0.31 ***
Higher Technical College	0.27	0.22***	0.25	0.19***	0.28	0.25 *
University or College Degree	0.24	0.37***	0.24	0.36***	0.23	0.39 ***
Net HH Income in KEUR	3.02	2.60***	3.02	2.70***	3.02	2.51 ***
Occupational Characteristics						
Labor Market Status						
Employed	0.77	0.72***	0.80	0.73***	0.74	0.71 **
Self-Employed	0.08	0.07	0.10	0.09	0.06	0.06
Unemployed	0.07	0.09***	0.07	0.09**	0.07	0.10 ***
Not Working	0.06	0.04***	0.01	0.02***	0.11	0.07 ***
In Education	0.01	0.07***	0.01	0.08***	0.01	0.07 ***
Unemployment Experience	0.98	0.83***	0.92	0.88	1.05	0.79 ***
Tenure in Current or Last Job						
$<$ 3 Years	0.32	0.54***	0.27	0.50***	0.37	0.59 ***
3-9 Years	0.27	0.31***	0.26	0.33***	0.27	0.29
$\geq$ 10 Years	0.38	0.13***	0.44	0.16***	0.33	0.10 ***
High Skilled Worker	0.18	0.22***	0.26	0.27	0.11	0.18 ***
Occupation Position						
Blue-Collar Worker	0.25	0.13***	0.34	0.21***	0.15	0.06 ***
White-Collar Worker	0.46	0.48	0.39	0.41	0.53	0.55
Civil Servent	0.06	0.08***	0.07	0.09**	0.05	0.07 *
Trainee	0.01	0.03***	0.01	0.03***	0.01	0.02 ***
Regional Characteristics						
Unemployment Rate	9.37	9.30	9.39	9.25	9.36	9.34
Gross Value Added	53.19	53.94***	53.10	53.76*	53.27	54.10 **
Population Density in 100	5.27	5.98***	5.22	5.66*	5.32	6.28 ***
Other Personality Variables						
Willingness to Take Risks[c]	4.60	5.04***	5.06	5.36***	4.16	4.73 ***
Openness[c]	4.48	4.69***	4.37	4.60***	4.58	4.79 ***
Conscientiousness[c]	5.92	5.78***	5.85	5.71***	5.98	5.85 ***
Extraversion[c]	4.84	4.90**	4.70	4.71	4.97	5.08 ***
Agreeableness[c]	5.35	5.35	5.19	5.23	5.51	5.46 *
Neuroticism[c]	3.84	3.81	3.59	3.53	4.08	4.07
Observations	107,636	1,598	52,353	788	55,283	810
Individuals	16,133	1,302	7,698	632	8,435	670

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: All numbers are shares unless stated otherwise. Stars give results from t-test for mean equivalence: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[a] Willingness to move has only 17,181 observations for 11,003 individuals.
[b] For this table, locus of control was adjusted to just have positive values by subtracting the lowest possible value.
[c] Willingness to Take Risks is measured on a Likert-Scale from 0 (= Risk-averse) to 10 (= Risk-prone) and the Big Five are measured with 3 items each on a Likert-Scale from 1 (= Does not apply at all) to 7 (= Applied to me perfectly). Factors are generated by averaging over the 3 items.

Table 3.A.4: Full Logit Estimation Results (Marginal Effects): Willingness to Move

	(1) All	(2) Men	(3) Women
	b/se	b/se	b/se
LOC Factor > Median	2.462***	3.440***	1.629*
	(0.695)	(1.006)	(0.957)
Socio-Economic Characteristics			
Age	-1.417***	-2.027***	-0.843
	(0.402)	(0.583)	(0.556)
Squared Age	0.017***	0.024***	0.009
	(0.005)	(0.007)	(0.007)
Female	-0.280		
	(0.851)		
Married	-3.135***	-2.765**	-4.041***
	(0.858)	(1.245)	(1.134)
German	-2.070	0.217	-4.225**
	(1.449)	(2.006)	(1.916)
Number of Children in HH	-0.445	0.620	-1.540**
	(0.424)	(0.598)	(0.608)
Disabled	1.020	0.894	1.145
	(1.444)	(1.956)	(2.030)
Home-Owner	-9.539***	-10.474***	-8.384***
	(0.778)	(1.090)	(1.078)
Net Household Income in KEUR	0.413*	-0.195	0.900***
	(0.224)	(0.331)	(0.319)
Highest School Degree (Ref: Lower or Intermediary School)			
No Degree	-0.597	1.909	-3.083
	(2.326)	(3.280)	(3.391)
Highschool Degree	4.965***	2.494	6.766***
	(1.055)	(1.538)	(1.308)
Vocational Education (Ref: Non)			
Apprenticeship	0.147	0.764	-0.496
	(0.937)	(1.351)	(1.308)
Higher Technical College	1.387	1.856	0.884
	(1.028)	(1.530)	(1.357)
University or College Degree	3.286***	2.461	3.615**
	(1.166)	(1.689)	(1.485)
Occupational Characteristics			
Unemployment Experience	0.196	-0.039	0.297
	(0.186)	(0.289)	(0.244)
High Skilled Worker	3.379***	4.801***	1.345
	(1.086)	(1.364)	(1.622)
Labor Force Status (Ref: Employed)			
Self-Employed	4.098**	3.885*	5.367**
	(1.611)	(2.010)	(2.269)
Unemployed	5.493***	3.740	7.264***
	(1.814)	(2.431)	(2.263)
Not Working	2.841	12.131***	1.910
	(1.853)	(4.476)	(2.108)
In Education	22.817***	26.792***	13.578***
	(3.643)	(4.141)	(3.972)
Tenure (Ref: 0-2 Years)			
3-9 Years	-2.062**	-1.452	-2.733**
	(0.807)	(1.234)	(1.116)
$\geq$ 10 Years	-4.545***	-4.537***	-5.057***
	(0.920)	(1.365)	(1.318)
Tenure not available	-3.573*	-16.104***	1.128
	(2.099)	(4.783)	(2.664)
Occupation Position (Ref: Blue-Collar)			
White-Collar Worker	2.259**	3.142**	1.911
	(1.044)	(1.434)	(1.608)
Civil Servent	1.489	3.480	-1.798

Continued on next page...

... continued from previous page

	(1) All	(2) Men	(3) Women
	(1.919)	(2.562)	(2.769)
Trainee	4.693	5.568	3.069
	(3.895)	(5.367)	(4.775)
NACE Industry (Ref: Manufacturing)			
Agriculture	-5.211*	-8.248*	-0.910
	(3.009)	(4.714)	(5.484)
Mining, Quarrying, Energy, Water	-2.766	-1.762	-6.955
	(3.088)	(3.810)	(7.955)
Chemicals, Pulp, Paper	3.033*	3.548	0.693
	(1.844)	(2.297)	(2.911)
Construction	-1.309	-0.489	-2.990
	(1.680)	(2.024)	(3.777)
Iron/Steel	-2.183	-1.758	-4.844
	(1.881)	(2.366)	(3.983)
Textile/Apparel	-2.696	-1.806	-5.467
	(3.965)	(8.178)	(5.205)
Wholesale/Retail	0.244	-0.521	-1.241
	(1.440)	(2.076)	(2.263)
Transport/Communication	1.763	2.933	-1.156
	(1.903)	(2.333)	(3.229)
Public Service	1.748	4.598***	-1.885
	(1.304)	(1.778)	(2.101)
Financial/Private Services	1.714	3.980**	-1.738
	(1.418)	(1.893)	(2.262)
NACE Industry Not Categorized	3.038*	6.676***	-1.649
	(1.677)	(2.345)	(2.463)
Regional Characteristics			
Unemployment Rate	0.059	0.192	-0.044
	(0.163)	(0.236)	(0.223)
Gross Value Added	0.002	0.045	-0.042
	(0.056)	(0.080)	(0.077)
Population Density in 100	0.165***	0.203***	0.117
	(0.053)	(0.077)	(0.072)
East-Germany	-10.911***	-14.912***	-8.649***
	(1.241)	(2.098)	(1.988)
Other Personality Variables			
Will. to Take Risks (std.)	3.395***	3.895***	2.591***
	(0.404)	(0.554)	(0.543)
Openness (std.)	1.469***	1.254**	1.575***
	(0.413)	(0.594)	(0.568)
Conscientiousness (std.)	-0.461	-0.903*	0.136
	(0.378)	(0.544)	(0.523)
Extraversion (std.)	-1.467***	-1.906***	-1.068*
	(0.417)	(0.591)	(0.575)
Agreeableness (std.)	-1.628***	-0.808	-2.398***
	(0.393)	(0.557)	(0.532)
Neuroticism (std.)	0.701*	1.220**	0.229
	(0.401)	(0.567)	(0.540)
Observations	17,181	8,362	8,819

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level.

Table 3.A.5: Full Logit Estimation Results (Marginal Effects): Internal Migration between t and t + 1

	(1) All	(2) Men	(3) Women
	b/se	b/se	b/se
LOC Factor > Median	0.189**	0.251**	0.126
	(0.083)	(0.127)	(0.111)
Socio-Economic Characteristics			
Age	-0.105**	-0.112	-0.074
	(0.049)	(0.072)	(0.066)
Squared Age	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
Female	-0.016		
	(0.099)		
Married	-0.626***	-0.453***	-0.833***
	(0.098)	(0.154)	(0.147)
German	0.346**	0.353	0.420
	(0.145)	(0.251)	(0.266)
Number of Children in HH	-0.319***	-0.270***	-0.392***
	(0.063)	(0.088)	(0.091)
Disabled	-0.102	-0.349	0.099
	(0.207)	(0.322)	(0.293)
Home-Owner	-1.140***	-1.319***	-1.268***
	(0.083)	(0.150)	(0.156)
Net Household Income in KEUR	-0.028	0.009	-0.049
	(0.028)	(0.030)	(0.044)
Highest School Degree (Ref: Lower or Intermediary School)			
No Degree	-0.107	-0.153	-0.047
	(0.347)	(0.508)	(0.551)
Highschool Degree	0.526***	0.391**	0.569***
	(0.115)	(0.172)	(0.145)
Vocational Education (Ref: Non)			
Apprenticeship	-0.115	0.047	-0.297**
	(0.102)	(0.141)	(0.151)
Higher Technical College	-0.014	-0.055	0.015
	(0.117)	(0.178)	(0.159)
University or College Degree	0.477***	0.490***	0.375**
	(0.132)	(0.180)	(0.158)
Occupational Characteristics			
Unemployment Experience	-0.082**	-0.035	-0.135***
	(0.033)	(0.051)	(0.040)
High Skilled Worker	0.101	0.039	0.175
	(0.120)	(0.162)	(0.166)
Labor Force Status (Ref: Employed)			
Self-Employed	0.262	0.136	0.582*
	(0.231)	(0.278)	(0.317)
Unemployed	1.343***	0.766***	1.430***
	(0.311)	(0.255)	(0.283)
Not Working	0.741**	1.370***	0.823***
	(0.322)	(0.402)	(0.309)
In Education	1.778***	1.389***	1.258***
	(0.408)	(0.277)	(0.320)
Tenure (Ref: 0-2 Years)			
3-9 Years	-0.150*	-0.095	-0.205*
	(0.084)	(0.125)	(0.120)
≥ 10 Years	-0.800***	-0.845***	-1.161***
	(0.098)	(0.193)	(0.210)
Tenure not available	-0.191	-0.115	-0.218
	(0.205)	(0.383)	(0.288)
Occupation Position (Ref: Blue-Collar)			
White-Collar Worker	0.274**	0.140	0.575**
	(0.138)	(0.180)	(0.235)

Continued on next page...

... continued from previous page

	(1) All	(2) Men	(3) Women
Civil Servent	0.861***	0.782***	0.717**
	(0.313)	(0.285)	(0.329)
Trainee	1.468***	1.235***	1.028**
	(0.491)	(0.353)	(0.407)
NACE Industry (Ref: Manufacturing)			
Agriculture	0.139	-0.157	0.437
	(0.403)	(0.469)	(0.530)
Mining, Quarrying, Energy, Water	-0.122	-0.034	-0.412
	(0.383)	(0.499)	(0.570)
Chemicals, Pulp, Paper	-0.077	-0.088	-0.199
	(0.235)	(0.292)	(0.363)
Construction	-0.372*	-0.517*	0.186
	(0.207)	(0.271)	(0.442)
Iron/Steel	-0.396*	-0.309	-1.390*
	(0.227)	(0.288)	(0.817)
Textile/Apparel	0.330	-0.145	0.223
	(0.538)	(1.024)	(0.477)
Wholesale/Retail	0.097	0.296	-0.157
	(0.187)	(0.237)	(0.261)
Transport/Communication	-0.134	0.112	-0.632
	(0.237)	(0.288)	(0.406)
Public Service	-0.136	0.127	-0.450*
	(0.160)	(0.199)	(0.247)
Financial/Private Services	-0.117	0.014	-0.315
	(0.168)	(0.214)	(0.258)
NACE Industry Not Categorized	-0.145	-0.179	-0.265
	(0.194)	(0.271)	(0.289)
Regional Characteristics			
Unemployment Rate	0.002	-0.014	0.021
	(0.017)	(0.025)	(0.023)
Gross Value Added	-0.019***	-0.023***	-0.013
	(0.006)	(0.008)	(0.008)
Population Density in 100	-0.001	-0.002	-0.002
	(0.005)	(0.008)	(0.007)
East-Germany	-0.557***	-0.689***	-0.493**
	(0.130)	(0.229)	(0.202)
Other Personality Variables			
Will. to Take Risks (std.)	0.240***	0.134*	0.303***
	(0.051)	(0.070)	(0.068)
Openness (std.)	0.089*	0.183***	-0.005
	(0.047)	(0.067)	(0.066)
Conscientiousness (std.)	0.036	0.058	0.004
	(0.042)	(0.062)	(0.058)
Extraversion (std.)	-0.103**	-0.195***	0.000
	(0.044)	(0.061)	(0.062)
Agreeableness (std.)	-0.010	-0.014	0.007
	(0.045)	(0.066)	(0.059)
Neuroticism (std.)	0.054	0.051	0.060
	(0.044)	(0.063)	(0.058)
Observations	109,234	53,141	56,093

Source: Socio-Economic Panel (SOEP), waves 1999-2015, version 33, doi:10.5684/soep.v33, own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level.

3.B Supplementary Anaylsis with IZA Evaluation Dataset

This section contains a detailed discussion of our ancillary analysis on geographic search range using the IZA Evaluation Dataset. We provide information about the data used, variable measurement, estimation strategy as well as descriptive statistics and full estimation results.

3.B.1 Data

In order to shed light one of the key predictions of our theoretical framework, we conduct an ancillary analysis of the relationship between locus of control and individuals' geographic search range. We are interested in knowing whether there is evidence that those job seekers with an internal locus of control search in across a broader geographic area. The SOEP data do not provide us with sufficient information on individuals' job-search to conduct this analysis; instead we make use of another German dataset – the IZA Evaluation Dataset. The IZA Evaluation Dataset targets a sample of Germans entering unemployment between June 2007 and May 2008. A 9 percent random sample from the monthly unemployment inflows of approximately 206,000 individuals identified in the administrative records was selected for interview. In wave 1, 17,396 interviews were completed with job seekers who had begun an unemployment spell approximately two months earlier. The data are ideal for our purposes because individuals are interviewed shortly after they become unemployed and are asked detailed questions about their job search and personality traits, including locus of control. We restrict our estimation sample to job seekers who are 25 to 55 years old to be consistent with our main SOEP analysis. We include observations from wave 1 as well as from waves 2 and 3 for those individuals who are still unemployed and searching for a job. This leaves us with in total 7,915 observations of which 6,987 are from wave 1 and an additional 499 (429) observations from wave 2 (wave 3).

Locus of Control In the IZA data, locus of control is measured using the same set of 10 items and on the same Likert scale as in the SOEP. We therefore construct a unidimensional continuous locus of control index in an identical manner as that in our SOEP analysis (see Section 3.4.2 for more details). An overview of the descriptive statistics of the LOC items in the data can be found in Table 3.B.1, while Figure 3.B.1 (upper panel) depicts the distribution of the continuous variable.

Geographic Search Range The IZA data also contain a measure of whether survey participants have applied for jobs for which they would have to make a residential move. Job seekers who answer this question with "yes" are considered to engage in "geographically distant" job search. For those who respond that they have applied for jobs that require a residential move, we also know the maximum search distance (in km) and, for a subset of individuals, whether migration would involve moving across state, national or continental boundaries. Table 3.B.2

provides a brief overview of the differences in the geographic range of internal and external job seekers.

Standard mean equality tests indicate that the internal job seekers are significantly more likely to indicate that they have searched for a job that would require them to change residences. This pattern is also supported by the kernel densities computed separately for movers and non-movers in Figure 3.B.1 (bottom panel). The maximum application distance is also significantly higher for internals (921 km) than for externals (753 km). There is an average search distance of 234 km for internals and 130 km for externals if all local (non-distant) searchers are set to zero. Internal job seekers engaged in distant job search are more likely to search Germany-wide, their external counter parts are more likely to restrict their search to their region or federal state. Although internal job seekers are significantly more likely to say that they are willing to change their residence for a job, there is no significant relationship between the willingness to accept long journeys to work and locus of control. This is consistent with our hypothesis that the relationship between locus of control and migration is mainly driven by search intensity and not by differences in the migration costs of internal vs. external job seekers.

3.B.2 Estimation Strategy

The choice of control variables was made to match the set of variables in our main estimation as close as possible. Restrictions due to data sensitivity imply that we lack information on regional characteristics, detailed household income, and an indicator of disability. Unfortunately, information on individual's willingness to take risk is collected only for a small subset of respondents and can thus not be included in the main analysis. Our labor market controls include the last occupation before unemployment as well as earlier unemployment experiences. The Big Five are surveyed with a reduced set of items in the IZA data and the construction of these traits is based on this reduced set. Information on home-ownership, highest school degree, highest vocational degree and unemployment experience is only available in wave 1 and is therefore imputed to waves 2 and 3. In order to identify potential non-linearities in the effects of LOC on distant search probabilities indicators for the quartiles of the LOC distribution are included as explanatory variables as an alternative to the continuous measure. Table 3.B.3 provides the key descriptive statistics for the estimation sample.

3.B.3 Results

The results of the main estimation can be found in Table 3.B.4. For the full estimation results corresponding to the estimations reported in column (7) see Table 3.B.5. Consistent with the predictions of our theoretical framework, internal job seekers are significantly more likely to engage in geographically distant job search. A one standard deviation increase in the extent to which someone is internal increases the probability of distant job search by 1.2

percentage points on average in the full specification (Column 7). This effect is largely driven by those individuals who are highly internal. Having a locus of control in the highest quartile increases the probability of geographically distant search by 3.34 percentage points relative to having a locus of control in the lowest quartile (Column 9). The baseline probability of geographically distant search is 22 percent, implying that these effects translate into an increase in geographically distant search of about 15 percent. Interestingly, these effects are stronger for women than for men which is consistent with previous evidence that the relationship between locus of control and labor market outcomes is often gender-specific (see Cobb-Clark 2015).

3.B.4 Tables and Figures

Table 3.B.1: Components of Locus of Control - IZA Eval. Dataset

Variable	mean
Components of locus of control (Scale: 1-7)[a]	
I1: How my life goes depends on me	6.03
I2: Compared to other people, I have not achieved what I deserve (R)	3.65
I3: What a person achieves in life is above all a question of fate or luck (R)	3.46
I4: If one is soc. or polit. active, one can have an effect on social conditions[b]	3.90
I5: I freq. have the experience that others have a controlling influence over my life (R)	2.88
I6: One has to work hard in order to succeed	6.19
I7: If I run up against difficulties in life, I often doubt my own abilities (R)	3.39
I8: Opportunities I have in life are determined by the social conditions (R)	4.54
I9: Inborn abilities are more important than any efforts one can make[b]	5.22
I10: I have little control over the things that happen in my life (R)	2.71
Observations	7,626

Source: IZA Evaluation dataset own calculations.
Notes: Items marked with a (R) are reversed prior to factor analysis in order to indicate an internal locus of control for high values.
[a] The LOC was surveyed on a 7-point likert scale from 1 for "I Disagree Completely" to 7 for "Agree completely".
[b] Items 4 and 9 are not included into the prediction of the latent factor.

Table 3.B.2: Descriptives of Outcome Variables by Locus of Control - IZA Eval. Dataset

	All	Internals	Externals
Distant Search (Dum)	0.22	0.26	0.18 ***
	[7,663]	[3,831]	[3,832]
Maximal Search Distance in km (incl. 0 km)[a]	182.17	234.02	130.33 ***
	[1,625]	[966]	[659]
Maximal Search Distance in km (if any)	853.00	921.54	752.54 **
	[7,609]	[3,804]	[3805]
Distant Search Intensity			
Own Town, Region	0.02	0.01	0.03 **
In Federal State	0.17	0.13	0.21 ***
In Germany	0.55	0.58	0.50 ***
In Europe	0.21	0.21	0.21
Outside Europe	0.06	0.06	0.06
	[1,677]	[993]	[684]
Concessions willing to take			
Long Journey to Work			
No	0.09	0.08	0.10
Maybe	0.46	0.45	0.46
Yes	0.46	0.47	0.44
	[2,051]	[1,075]	[976]
Change of residence			
No	0.50	0.47	0.52 **
Maybe	0.34	0.34	0.35
Yes	0.16	0.19	0.13 ***
	[2,048]	[1,073]	[975]

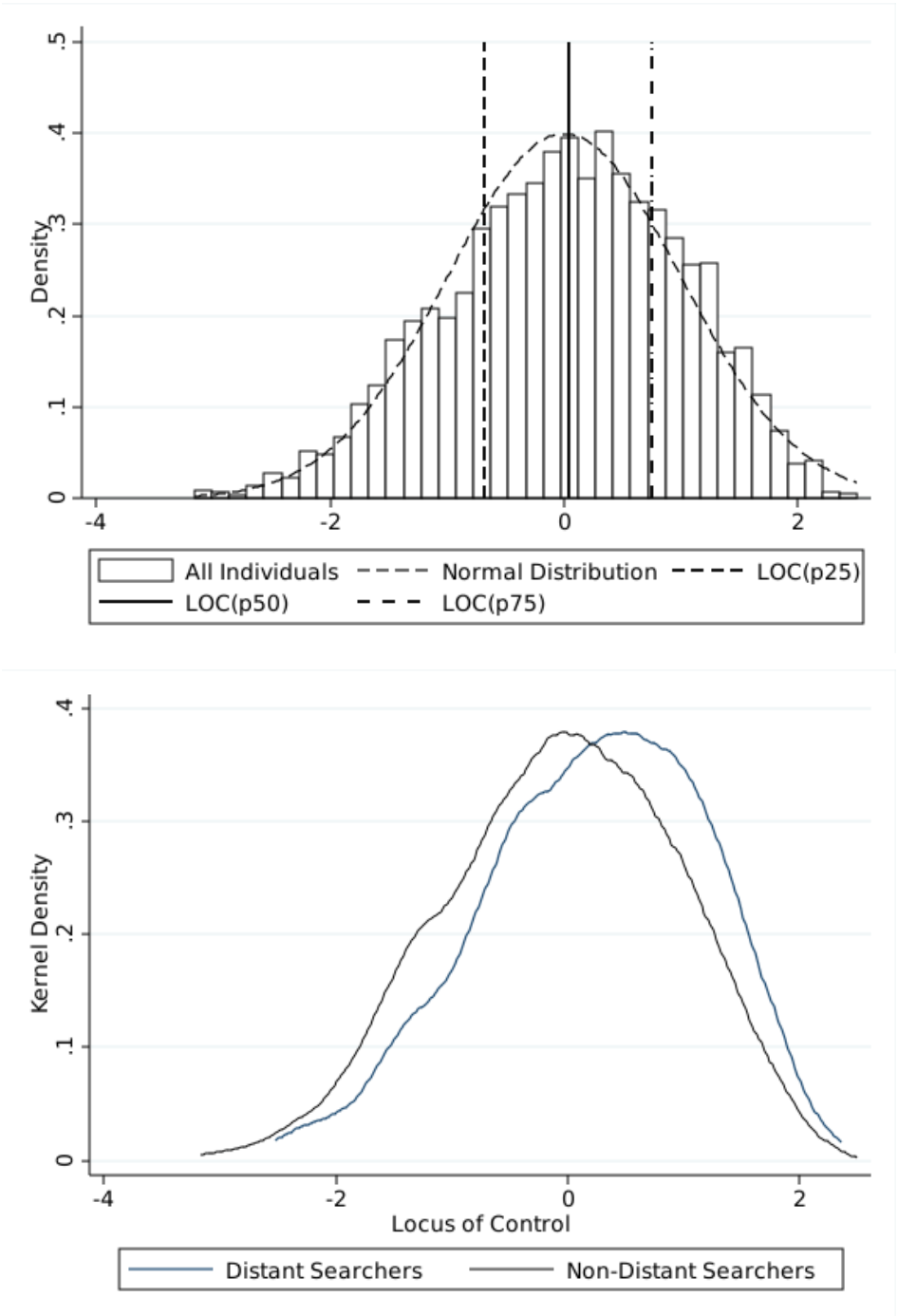
Source: IZA Evaluation Dataset, own calculations.
Notes: All numbers are shares unless stated otherwise. Number of observations in square brackets. Significance stars in the last column refer to significance level of t-test for mean equivalence between Externals and Internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[a] Distance is coded to 0 if individual indicates that she did not apply for a job for which she would have to move.

Table 3.B.3: Selected Descriptive Statistics by Internal Migration Status - IZA Eval. Dataset

	All		Male		Female	
	Non-Distant Seeker	Distant Seeker	Non-Distant Seeker	Distant Seeker	Non-Distant Seeker	Distant Seeker
Locus of Control (std.)[a]	3.10	3.38***	3.07	3.34***	3.14	3.40***
Locus of Control > Median	0.47	0.59***	0.46	0.58***	0.49	0.60***
Socio-Economic Characteristics						
Age	39.65	36.19***	39.88	34.60***	39.38	37.12***
Female	0.46	0.63***				
German	0.95	0.95	0.95	0.95	0.94	0.95
Married or living together	0.53	0.30***	0.58	0.22***	0.48	0.34***
Number of Children in HH	0.59	0.31***	0.69	0.23***	0.48	0.3***6
East-Germany	0.31	0.30	0.29	0.32	0.34	0.29***
Home-Owner	0.39	0.22***	0.41	0.20***	0.36	0.22***
Monthly Net-Income of the HH						
less or equal 1000 EUR	0.23	0.33***	0.20	0.36***	0.27	0.32***
1001 - 2000 EUR	0.41	0.41	0.38	0.40	0.45	0.41*
2001 - 3000 EUR	0.28	0.19***	0.32	0.18***	0.24	0.20**
3001 - 4000 EUR	0.08	0.07	0.11	0.06***	0.05	0.07**
Highest School Degree						
No Degree	0.01	0.01**	0.01	0.00	0.01	0.01**
Lower Secondary School	0.30	0.19***	0.24	0.10***	0.38	0.24***
Intermediate School	0.45	0.32***	0.50	0.34***	0.40	0.30***
Highschool	0.22	0.47***	0.25	0.55***	0.19	0.43***
Other	0.01	0.01	0.01	0.00*	0.01	0.02
Vocational or Higher Education						
No Vocational Education	0.09	0.05***	0.09	0.04***	0.08	0.05***
Apprenticeship	0.59	0.39***	0.56	0.34***	0.62	0.41***
Higher Technical College	0.15	0.16	0.15	0.12**	0.15	0.18**
University Degree	0.13	0.37***	0.15	0.45***	0.11	0.32***
Other Vocational Degree	0.04	0.04	0.04	0.05	0.04	0.03
Last Employment Status before Unemployment						
Regularly Employed	0.73	0.71**	0.68	0.65	0.80	0.74***
Subsidized Employment	0.07	0.06	0.06	0.05	0.07	0.07
Education, Milit. Serv.	0.04	0.15***	0.04	0.19***	0.04	0.12***
Parental Leave etc.	0.07	0.01***	0.12	0.02***	0.00	0.00
Unemployment Experience	2.69	2.79	2.36	2.57	3.07	2.92
Other Personality Variables						
Openness[d]	4.94	5.26***	4.92	5.33***	4.97	5.22***
Conscientiousness[d]	6.25	6.26	6.33	6.40**	6.17	6.18
Extraversion[d]	5.09	5.20***	5.18	5.36***	4.99	5.11***
Neuroticism[d]	3.83	3.55***	3.98	3.80***	3.65	3.40***
Observations	5,984	1,679	3,214	622	2,770	1,057

Source: IZA Evaluation dataset own calculations.
Notes: All numbers are shares unless stated otherwise.
Stars give results from t-test for mean equivalence: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[b] For detailed information on the measurement of locus of control see Section 3.4.2. For this table, locus of control was adjusted to just have positive values by subtracting the lowest possible value.
[d] Big Five are measured with a reduced set of items available from the data. See Section 3.B.2 for more details.

Figure 3.B.1: Distribution of Locus of Control in IZA Eval. Dataset



Source: IZA Evaluation dataset, own illustration.
Notes: Bottom panel provides the graphical illustration of the kernel density estimates using the Epanechnikov kernel function separately for distant searchers and non-distant searchers.

Table 3.B.4: Logit Estimation Results (Marginal Effects): Geographically Distant Search

	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se	(7) b/se	(8) b/se	(9) b/se
All									
LOC Factor	4.693***			1.725***			1.195**		
	(0.488)			(0.468)			(0.500)		
LOC > Med.		7.822***			2.420***			1.386	
		(0.952)			(0.912)			(0.949)	
Locus of Quartiles (Ref.: $[LOC_{min}, LOC_{P25}]$)									
$(LOC_{P25}, LOC_{P50}]$			4.043***			1.657			1.609
			(1.243)			(1.249)			(1.269)
$(LOC_{P50}, LOC_{P75}]$			7.040***			1.846			1.412
			(1.289)			(1.279)			(1.305)
$(LOC_{P75}, LOC_{max}]$			12.651***			4.798***			3.344**
			(1.348)			(1.305)			(1.383)
Observations	7,663	7,663	7,663	7,663	7,663	7,663	7,663	7,663	7,663
Pseudo R^2	0.014	0.011	0.014	0.156	0.155	0.156	0.164	0.163	0.164
Men									
LOC Factor	3.667***			1.205**			0.775		
	(0.611)			(0.580)			(0.629)		
LOC > Med.		6.117***			1.458			0.620	
		(1.199)			(1.120)			(1.175)	
Locus of Quartiles (Ref.: $[LOC_{min}, LOC_{P25}]$)									
$(LOC_{P25}, LOC_{P50}]$			4.474***			2.684*			2.844*
			(1.536)			(1.560)			(1.588)
$(LOC_{P50}, LOC_{P75}]$			7.514***			2.471			2.265
			(1.617)			(1.573)			(1.626)
$(LOC_{P75}, LOC_{max}]$			9.210***			3.381**			2.235
			(1.663)			(1.609)			(1.732)
Observations	3,828	3,828	3,828	3,828	3,828	3,828	3,828	3,828	3,828
Pseudo R^2	0.013	0.010	0.013	0.225	0.224	0.225	0.233	0.233	0.234
Women									
LOC Factor	5.118***			2.212***			1.545**		
	(0.748)			(0.724)			(0.774)		
LOC > Med.		8.022***			2.524*			1.126	
		(1.455)			(1.431)			(1.477)	
Locus of Quartiles (Ref.: $[LOC_{min}, LOC_{P25}]$)									
$(LOC_{P25}, LOC_{P50}]$			4.271**			2.217			1.933
			(1.948)			(1.937)			(1.974)
$(LOC_{P50}, LOC_{P75}]$			5.991***			1.309			0.588
			(1.993)			(1.987)			(2.018)
$(LOC_{P75}, LOC_{max}]$			14.313***			6.104***			4.160*
			(2.076)			(2.039)			(2.175)
Observations	3,827	3,827	3,827	3,827	3,827	3,827	3,827	3,827	3,827
Pseudo R^2	0.015	0.011	0.015	0.106	0.105	0.107	0.114	0.113	0.114
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Socio-Econ.				✓	✓	✓	✓	✓	✓
Personality							✓	✓	✓

Source: IZA Evaluation dataset own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level.
Full estimation results for specifications (7) can be found in Table 3.B.5. The full estimation results for specifications (1) - (6) as well as (8) and (9) can be obtained from the authors.

Table 3.B.5: Full Logit Estimation Results (Marginal Effects): Geographically Distant Search

	(1) All	(2) Men	(3) Women
	b/se	b/se	b/se
LOC Factor (std.)	1.195**	0.775	1.545**
	(0.500)	(0.629)	(0.774)
Socio-Demographic Characteristics			
Age	-1.339**	-0.909	-1.036
	(0.538)	(0.679)	(0.835)
Squared Age	0.013*	0.006	0.010
	(0.007)	(0.009)	(0.011)
Female	9.713***		
	(0.933)		
German	1.718	-0.512	3.557
	(1.882)	(2.423)	(2.935)
Married	-5.716***	-8.465***	-4.600***
	(1.080)	(1.347)	(1.738)
Number of Children in Household	-2.868***	-6.679***	-0.637
	(0.753)	(1.058)	(1.169)
East-Germany	-0.985	1.378	-2.333
	(0.989)	(1.235)	(1.545)
Home-Owner	-5.739***	-1.592	-8.830***
	(0.996)	(1.283)	(1.511)
Monthly Net-income of the HH (Ref: less or equal 1000 EUR)			
1001 - 2000 EUR	-1.520	-1.051	-1.648
	(1.079)	(1.268)	(1.711)
2001 - 3000 EUR	-4.648***	-5.089***	-3.516*
	(1.282)	(1.509)	(2.059)
3001 - 4000 EUR	-4.331**	-5.160***	-0.625
	(1.701)	(1.893)	(3.067)
Highest School Degree (Ref: Non)			
Lower Secondary School	3.879	1.752	4.546
	(5.338)	(7.504)	(7.904)
Intermediate School	6.117	4.851	6.011
	(5.169)	(6.993)	(7.932)
Highschool	14.083**	8.658	16.416*
	(6.347)	(8.673)	(9.574)
Other School	14.301*	-4.038	20.693*
	(8.349)	(11.351)	(11.432)
Vocational or Higher Education (ref: Non)			
Apprenticeship	3.879**	1.453	4.781*
	(1.889)	(2.516)	(2.802)
Higher Technical College	12.000***	3.539	17.149***
	(2.514)	(3.068)	(3.688)
University or College Degree	23.066***	14.083***	28.694***
	(2.901)	(3.874)	(4.198)
Other Vocational Degree	8.740**	7.115*	7.181
	(3.424)	(4.229)	(5.139)
Last Employment Status before Unemployment (Ref: Reg. Employment)			
Subsidized Employment	1.170	1.146	1.238
	(1.852)	(2.260)	(2.811)
Education, Training, Military Serv.	7.660***	7.901***	6.380**
	(2.034)	(2.376)	(3.195)
Parental Leave etc.	-12.128***	-5.895**	-2.735
	(2.100)	(2.366)	(9.953)
Ealier Unemployment Experience in Months	0.085	0.241	-0.095
	(0.120)	(0.150)	(0.183)
Other Personality Variables			
Openness	2.460***	1.830***	2.904***
	(0.401)	(0.499)	(0.617)
Conscientiousness	0.592	1.162	0.045
	(0.586)	(0.790)	(0.872)

Continued on next page...

... continued from previous page

	(1) All	(2) Men	(3) Women
Extraversion	0.213	0.381	0.246
	(0.443)	(0.551)	(0.681)
Neuroticism	-1.052**	-0.642	-1.497**
	(0.411)	(0.510)	(0.638)
Observations	7,663	3,828	3,827

Source: IZA Evaluation Dataset, own calculations.
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Marginal effects are reported in percentage points. Standard Errors in parentheses. Standard Errors are clustered on person-level.

CHAPTER 4

Drinking is Different! Examining the Role of Locus of Control for Alcohol Consumption

*This chapter is based on joint work with **Marco Caliendo**.*

4.1 Introduction

Unlike other practices of unhealthy behavior, such as smoking, physical inactivity, unhealthy diets, substance use or risky sexual behavior, alcohol consumption is still largely accepted within most societies. According to data from the World Health Organisation 71.7 percent of US-Americans and 79.4 percent of Germans over the age of 15 consumed alcohol in 2016 while the rate of lifetime abstainers was with 9.2 percent (USA) and 7.9 percent (Germany) vanishingly low in the same year. Even heavy episodic drinking is a surprisingly common behavior in both countries. In Germany 43.1 percent and in the USA 36.4 percent of drinkers engaged in a heavy drinking episode in the past 30 days (World Health Organization, 2016).[1] At the same time, a huge body of studies in multiple fields such as medicine as well as all social science, including economics, have empirically identified severe negative effects of heavy alcohol consumption, alcohol abuse and addiction on mental and physical health (see e.g. Grønbæk, 2009; Boffetta and Hashibe, 2006; Thakker, 1998; Merikangas *et al.*, 1998; Keil *et al.*, 1998; Corrao *et al.*, 2004) as well as social and economic outcomes (see e.g. Mullahy and Sindelar, 1993, 1996; Jones and Richmond, 2006; Macdonald and Shields, 2004). Thus, the consequences are inextricably linked to considerable microeconomic and macroeconomic costs, e.g. through the strain it puts on individual health care expenditures and labor market perspectives.

As opposed to this, a growing empirical literature finds positive effects of moderate drinking on health, such as e.g. mortality, heart disease, cancer as well as dementia (see e.g. Rehm *et al.*, 2001; Grønbæk *et al.*, 2000; Renaud *et al.*, 1998; Ronksley *et al.*, 2011; Peters *et al.*, 2008), and on economic outcomes (see e.g. Macdonald and Shields, 2001; Peters and Stringham, 2006; Ours, 2004; French and Zarkin, 1995; Hamilton and Hamilton, 1997). Not least because of this ambiguity in the literature about the consequences of different levels of alcohol consumption on health as well as economic and social outcomes, decision making with respect to drinking is associated with a high level of uncertainty about positive and negative future consequences. This is why individual subjective beliefs about risks and personal control are likely to play a crucial role when explaining individual differences in drinking behavior as an important health decision.

Based on these considerations, this paper attempts to investigate the role of the personality trait locus of control (LOC) on individual alcohol consumption. Locus of control can be characterized as a *"generalized attitude, belief, or expectancy regarding the nature of the causal relationship between one's own behavior and its consequences"* (Rotter, 1966) and describes whether individuals believe in the effects of their own actions on their lives' future outcomes. While an individual with an internal LOC believes that she is in control of the consequences of her own actions, an external individual attributes lives outcomes to luck, chance, fate or other

[1] According to the World Health Organization (2016), heavy episodic drinking is defined as the consumption of at least 60 grams of pure alcohol which corresponds to approximately 6 standard alcoholic drinks.

external forces. LOC has already been shown to have an important effect on behavior and decision making in such areas as human capital investment (Coleman and DeLeire, 2003), job search effort (McGee and McGee, 2016; Caliendo *et al.*, 2015b), savings (Cobb-Clark *et al.*, 2016), occupational attainment (Heywood *et al.*, 2017; Cobb-Clark and Tan, 2011), entrepreneurial activity (Caliendo *et al.*, 2014), labor market mobility (Caliendo *et al.*, 2015a), investment behavior (Salamanca *et al.*, 2016; Pinger *et al.*, 2018) and, most related to the paper at hand, health behavior such as smoking, healthy diets and exercise (Cobb-Clark *et al.*, 2014). Nevertheless, to the best of our knowledge, no paper has yet attempted to empirically identify an effect of LOC on drinking behavior in a well-conceived and detailed manner. Building on the existing empirical findings on the association between LOC and individual health-investments as well as risky behavior, which are summarized in detail in Section 4.2, our paper intends to sheet light on the role of LOC for drinking as a very important domain of risky health behavior. We are doing so by empirically investigating the relationship between an internal locus of control and self-reported levels of alcohol consumption and by additionally discussing and analysing the potential mechanism and channels behind it.

The estimations are conducted using the extensive information available from the Socio-Economic Panel (SOEP, 2017), a large representative longitudinal household panel from Germany. The SOEP not only includes detailed socio-economic information but also regularly surveys individuals' locus of control as well as their health behavior including alcohol consumption. Using this data, we estimate the effects of an individual's locus of control on her probability of indicating higher levels of alcohol consumption. We find a significant positive effects of having an internal LOC on the individual probability to be a moderate or heavy drinker as measured by occasional or regular consumption of at least one out of four beverage types. The effects also hold if we control for a large list of socio-economic, health-related and personality and preference controls and are of considerable economic significance as they are of similar magnitude as effects of traditional preference measures such as risk aversion and time preferences. In addition, the estimated effects are robust with respect to the definition of the LOC variable, to the use of a more objective measure of consumption as well as to controlling for compensating behavior in other health-domains as a potential confounder.

Based on these findings, we argue that two main mechanisms are likely to be responsible for this relationship. On the one hand, LOC is likely to be highly predictive for the amount of individual investment into social and occupational networks and thus also for drinking opportunities as the attendance of social gatherings is often inextricably linked with alcohol consumption. On the other hand, an increased perception of individual control (being internal) might decrease the perceived importance of risk for life's outcomes. The future risks of alcohol consumption might thus be underestimated if individuals believe in their own ability to cope with or prevent the negative consequences of risky behavior.

The outline of the paper is as follows. Section 4.2 gives a brief introduction into the earlier literature. Following this, Section 4.3 describes the data as well as the estimation strategy and Section 4.4 presents an overview of the results of the main estimation. Based on this, Section 4.5 argues about potential channels behind the estimated effects. Section 4.6 presents a number of robustness checks. The paper concludes in Section 4.7.

4.2 The Economics of Risky Health Behavior

Traditional Models Analysing the drivers of risky health-related behavior, such as alcohol consumption, smoking, unprotected sex, unbalanced diets as well as a lack of exercise, has a relatively long tradition in the modern economic literature. Motivated by the medical literature on the severe negative consequences of risky health behavior, economists have started analysing the drivers of these types of behavior. The standard economic approach to study risky health-related behavior concentrated on a traditional rational choice model assuming that individuals maximize the present discounted value of lifetime utility. In line with the well-known health-capital model by Grossman (1972, 2000), individuals are thus assumed to drink if the instantaneous pleasure from drinking now is higher than the expected utility from investing in the future health stock by abstaining or reducing consumption plus the monetary costs of purchasing the alcohol on the market. Especially relevant to the research on alcohol consumption and smoking, another strand of literature evolved over the years, which puts a specific focus on the economic rationale behind addictions. Based on earlier work from psychology (see e.g. Rachlin, 1997), the literature assumes that current levels of consumption are positively affected by past consumption levels and thus models addiction using autocorrelation within the demand functions for addictive goods (Becker and Murphy, 1988; Chaloupka, 1991). In order to explain medium levels of unhealthy behavior, especially non-standard approaches of time discounting, i.e. present-biased preferences as well as time inconsistent preferences have been discussed (see e.g. Cutler *et al.*, 2003; Gruber and Köszegi, 2001).

Behavioral Approaches Over the years, multiple other approaches of explaining risky health behavior have evolved. Peer effects have been found to be highly important[2] (see e.g. Manski, 2000; Christakis and Fowler, 2007), with the empirical literature especially showing strong peer dependencies in alcohol consumption for adolescents (see e.g. Duncan *et al.*, 2005; Argys and Rees, 2008; Lundborg, 2006). But also information constraints (see e.g. Kenkel, 1991; Viscusi, 1990) as well as bounded rationality and cognitive limitations such as through self-control failures in line with the seminal work by e.g. Simon (1982) and Thaler and Shefrin

[2] An important term in this context is the so called "bandwagon effect" which describes a situation in which individuals derive additional utility from the fact that they are consuming the same goods as peers (Leibenstein, 2006).

(1981) are discussed as important issues in this domain (see e.g. Suranovic *et al.*, 1999). In line with this, a large literature focused on the importance of education and cognitive ability for preventing bounded rationality and information constraints in the health context (see e.g. Farrell and Fuchs, 1982; Cutler and Lleras-Muney, 2010). A detailed discussion of the economics behind risky health behavior in general and alcohol consumption in specific as well as a profound literature review can be found e.g. in Cawley and Ruhm (2011) and Cook and Moore (2000).

Psychological Determinants While the standard economic models above would attribute unexplained individual heterogeneity in risky health behavior to idiosyncratic shocks, behavioral economics has motivated especially modern empirical studies to investigate the psychological black box behind such behavior. In the past years the literature on psychological determinants of risky health behavior in general and alcohol consumption in specific has been on the rise. Starting with the early psychological and medical literature, largely small-scale empirical studies have found relationships of personality traits such as susceptibility to peer pressure, self-esteem, extraversion and neuroticism on substance use and smoking in adolescence (Dielman *et al.*, 1987; Wijatkowski *et al.*, 1990). In line with this, the later psychological literature managed to observe important relationships of personality traits such as the Big Five and risky health behavior of adults (see e.g. Vollrath *et al.*, 1999; Booth-Kewley and Vickers, 1994; Lemos-Giráldez and Fidalgo-Aliste, 1997).

Locus of Control With respect to locus of control, the psychological literature largely focused on very specific measures of LOC which are directly linked to the perceived control in the health domain and found important relationships between health-related LOC and substance use in adolescence (Carman, 1974; Dielman *et al.*, 1987) as well as health behavior such as diets, exercise, smoking and seat belt use in adulthood (see e.g. Wallston *et al.*, 1978; Wallston and Wallston, 1978; Furnham and Greaves, 1994; Lemos-Giráldez and Fidalgo-Aliste, 1997; Holt *et al.*, 2014). In line with this, Steptoe and Wardler (2001) find that perceptions of high control through chance as well as powerful others are associated with higher alcohol consumption levels. As opposed to this large body of literature in psychology, the economic literature which attempts to incorporate locus of control into their health-behavior models is quite rare. Mendolia and Walker (2014a) for example analyse the effect of locus of control as well as self-esteem and work ethics on health behavior, as measured by alcohol consumption, drug use, unprotected and early sexual activity and sports, for a group of adolescents aged 15-16 years. With respect to alcohol consumption and locus of control, they only find a weak link of having an external perception of control with the frequency of getting drunk when drinking but no significant association with regular drinking. Chiteji (2010) analyses the role of future orientation and self-efficacy on drinking and exercising. They find that self-efficacy, which is strongly linked to locus of control, is negatively associated with drinking and positively

associated with exercising. Most closely related to our paper, Cobb-Clark *et al.* (2014) discuss the role of general locus of control for individual health behavior on multiple dimensions including diet, smoking and exercising using Australian survey data from the HILDA. They find positive effects of LOC on healthy habits such as healthy diets, non-smoking and regular exercises. Based on an extensive heterogeneity analysis, the authors find that for men the likely mechanism behind this relationship is a difference in the subjective expectations about returns to investments into health. Thus, internal men are likely to expect higher returns to them being health conscious and are thus more likely to invest in healthy habits. As opposed to this, women seem to have stronger preferences for health investments in general. In line with this research, we would also expect negative effects of LOC on alcohol consumption if drinking less or totally abstaining from alcohol is assumed to lead to positive outcomes for the own health in the future. Nevertheless, as a side note, Cobb-Clark *et al.* (2014) identify significantly positive effects of LOC on alcohol consumption in excessive amounts (i.e. binge drinking). The authors argue that an increased perception of control might be correlated with a higher belief about the ability to cope with the consequences of binge drinking and additionally alcohol might not be perceived as a behavior which strongly affects health outcomes in general. Nevertheless, their data does not allow them to analyse drinking behavior in more detail. This is where our paper is able to make a major contribution in order to complete the picture on locus of control and risky health behavior.

4.3 Data and Empirical Approach

Building upon the existing literature, we estimate the relationship between an internal locus of control and self-reported levels of alcohol consumption using German panel data. The estimations are conducted using the extensive information available from the Socio-Economic Panel (SOEP, 2017), a large representative longitudinal household panel from Germany. The SOEP not only includes detailed socio-economic information but also surveys individuals' locus of control as well as their health behavior, including alcohol consumption, on a regular basis. For more information on the SOEP see Goebel *et al.* (2018). Our sample consists of all observations for individuals between the age of 20 to 70 years for the waves 2006, 2008 and 2010 in which we observe the self-assessed amount of alcohol consumption.

Table 4.A.1 gives an overview over the main summary statistics for the sample. The final estimation sample consists of 36,134 observations for 15,482 individuals due to the panel nature of the data. Individuals are thus observed up to three times. 17,418 (7,478) of those observations (individuals) are men and 18,716 (8,004) are women. The later estimations will always be reported separately for men and women in order to take care of important gender specific effects which are common to the literature (see e.g. Cobb-Clark *et al.*, 2014).

4.3.1 Alcohol Consumption

In 2006, 2008 and 2010, individuals were asked to rate their consumption levels of four different types of alcoholic beverages (beer, wine, spirits, mixed drinks) on a scale from 1 (regularly) to 4 (never). Based on those answers and guided by the work of Ziebarth and Grabka (2009), we generate an ordinal measure combining all those answers. The variable categorizes individuals into the following four groups:

Abstainers - No consumption of all four types

Rare Drinkers - Seldom drinking for at least one type but no occasional consumption

Moderate Drinkers - Occasional drinking for at least one type but no regular consumption

Heavy Drinkers - Regular drinking for at least one type

In line with this categorization, the terms moderate and occasional as well as heavy and regular are going to be used interchangeably in the following. As is already noted by Ziebarth and Grabka (2009), the main drawback of this measurement is the rather vague and subjective character, as no concrete information about the exact quantity of alcohol consumption is collected. Nevertheless, we are able to test the reliability of this measure as well as the sensitivity of our results with respect to the subjectivity of the reported amounts using a new measure available in the SOEP wave 2016. We find that the subjectively reported levels are quite reliable based on a correlation of 0.634 between subjective levels in 2010 and concrete frequencies in 2016 which is fairly high given that the observations lie six years apart. For a more detailed discussion of the sensitivity of our results in this respect, see Section 4.6.2.

In a second step, we then generate binary indicators based on the defined groups in order to enable the analysis of non-parallel effect as well as the calculation of marginal effects. The following overview summarizes all four drinking indicators D:

Drinking Indicator	**Never + Rare**	**Moderate**	**Heavy**
D_1 Moderate + Heavy	$D = 0$	$D = 1$	$D = 1$
D_2 Moderate	$D = 0$	$D = 1$	missing
D_3 Heavy I *(vs. Moderate)*	missing	$D = 0$	$D = 1$
D_4 Heavy II *(vs. Never/Rare)*	$D = 0$	missing	$D = 1$

Our main outcome variable is the binary indicator for moderate or heavy drinkering (D_1), which captures all individuals who reports occasional or regular consumption on at least one of the beverage types. In order to identify non-parallel effects, three alternative indicators D_2 - D_4 are used in an additional step, which are all based on a selective sample of drinkers as either one of the three possible groups is set to missing.[3] These indicators enable us to make statements about the effects on the probabilities of moderate and heavy consumption

[3] Abstainers and rare drinkers are always pooled into one reference group for the main estimations.

Table 4.1: Descriptive Analysis - Alcohol Consumption

	All	**Women**	**Men**
Alcohol Consumption			
Abstainers	0.12	0.15	0.08
Rare Drinkers *(Seldom Consumption)*	0.29	0.36	0.21
Moderate Drinkers *(Occasional Consumption)*	0.42	0.40	0.44
Heavy Drinkers *(Regular Consumption)*	0.18	0.09	0.27
Moderate or Heavy Drinking D_1 (Binary Indicators)			
All	0.60	0.49	0.71
Beer	0.39	0.20	0.60
Wine	0.39	0.41	0.36
Spirits	0.11	0.07	0.15
Mixed Drinks	0.06	0.06	0.06
Observations	36,134	18,716	17,418

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.

independent from each other. Nevertheless, it has to be noted that the estimated effects using those outcome variables are at risk of being biased by sample selection. The estimated results thus have to be interpreted with care and are only ancillary evidence in addition to the main indicator.

Table 4.1 gives an overview of the shares of alcohol consumption in the sample. In the full sample, 60% of the individuals are counted as being moderate or heavy drinkers, as they indicate that they consume at least one beverage type occasionally (42%) or regularly (18%), and 12% can be characterized as abstainers (no alcohol consumption at all). In line with expectations, the share of drinkers is distinctly lower for women (40% moderate drinkers and 9% heavy drinkers) than for men (44% moderate drinkers and 27% heavy drinkers). While 15% of all women are abstainers, this is only true for 8% of men. Men and women also differ with respect to the beverage types they consume. While 60% (15%) of men consume beer (spirits) at least occasionally, only 20% (7%) of women do so. As opposed to this, the share of moderate wine-drinkers is higher for women (41% as compared to 36%).

4.3.2 Locus of Control

For our sample, locus of control is available in the years 2005 and 2010. In those waves, SOEP respondents were asked to assess their agreement with a list of 10 statements (items), all related to the degree of influence on life's outcomes. Responses were measured on a seven-point Likert scale ranging from 1 ('disagree completely') to 7 ('agree completely'). A list of the set of items can be found in Table 4.2.

In order to construct our locus of control variable, we conduct an exploratory factor analysis with which we investigate the way these items load onto latent factors.[4] The factor

[4] A loading plot pooled for all years can be found in Figure 4.B.1 in the Supplementary Material. The detailed results of this factor analysis are available from the authors on request.

Table 4.2: Components of Locus of Control (waves 2005 and 2010)

No	Item	*mean*	*SD*
Q:	The following statements apply to different attitudes towards life and the future. To what degree do you personally agree with the following statements? *Scale: 1 (Disagree completely) - 7 (Agree completely)*		
I1:	How my life goes depends on me	5.49	(1.29)
I2:	Compared to other people, I have not achieved what I deserve (-)	3.26	(1.77)
I3:	What a person achieves in life is above all a question of fate or luck (-)	3.51	(1.64)
I4:	If a person is socially [...] active, he/she can have an effect on social conditions[b]	3.71	(1.61)
I5:	I have the experience that others have a controlling influence over my life (-)	3.12	(1.68)
I6:	One has to work hard in order to succeed	5.95	(1.11)
I7:	If I run up against difficulties in life, I often doubt my own abilities (-)	3.26	(1.64)
I8:	The opportunities that I have in life are determined by the social conditions (-)	4.50	(1.47)
I9:	Inborn abilities are more important than any efforts one can make[b]	4.84	(1.33)
I10:	I have little control over the things that happen in my life (-)	2.67	(1.47)
	Observations	24,195[a]	

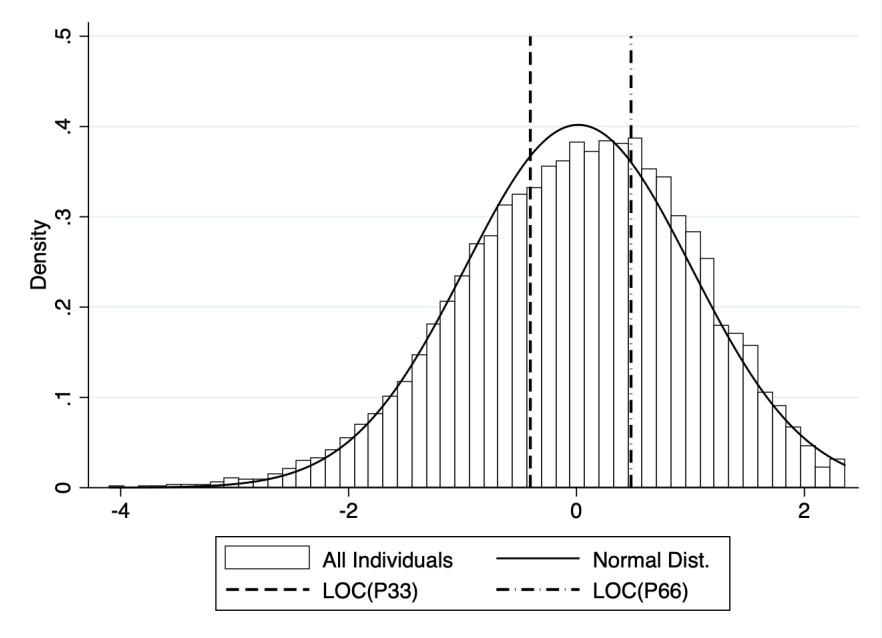
Source: SOEP, waves 2005 and 2010, version 33, doi:10.5684/soep.v33.
Notes: Items marked with a (-) are reversed prior to factor analysis.
[a] In this table, the item means and SD are computed for the observation waves 2005 and 2010 only, independent from the sample definition.
[b] Items 4 and 9 are not included into the final factor.

analysis reveals that items 1 and 6 clearly load onto the first factor – which is interpreted as internal LOC –, while items 2, 3, 5, 7, 8 and 10 clearly load onto the second factor – interpreted as external LOC. Items 4 and 9 are not included in the following due to ambiguous loadings.[5] Afterwards, we use a two-step process to create a continuous unidimensional locus of control variable, consistent with the previous literature (see e.g. Piatek and Pinger, 2016). Based on the exploratory factor analysis, we first reverse the scores for the external items (items 2, 3, 5, 7, 8 and 10) such that all eight items are increasing in internality.[6] Afterwards, we use confirmatory factor analysis to extract a single factor. This has the advantage of avoiding equal weighting of all item and instead allow the data to drive how each item is weighted in the overall index. As has been shown by Piatek and Pinger (2016), simple averaging risks measurement error and attenuation bias. The resulting factor is therefore increasing in internal locus of control and its distribution is shown in Figure 4.1.

There is evidence that LOC is relatively stable for the working-age population (see e.g. Preuss and Hennecke, 2018; Cobb-Clark and Schurer, 2013) and thus concerns with respect to reverse causality are rather small. With respect to alcohol consumption potential endogeneity concerns are additionally likely to apply only to rather extreme levels of consumption which are arguable very seldom also in our sample. In order to further minimize those concerns, we ensure that the LOC factor is never measured after the period in which we measure alcohol consumption. Thus, the information is then imputed forwards into the years in which we observe alcohol consumption, i.e. the LOC from 2005 is used as the explanatory variable for

[5] Item 4 does not clearly load onto either of the two factors and Item 9 loads onto the internal factor but an intuitive attribution based on the item's wording would point more in the direction of an external item.
[6] These items are marked with an (-) in Table 4.2.

Figure 4.1: Distribution of Locus of Control



Source: SOEP, waves 2005 and 2010, version 33, doi:10.5684/soep.v33, own illustration.

alcohol consumption in 2006 and 2008 and LOC from 2010 is used for consumption in 2010. Additionally, we will test the sensitivity of the results with respect to the construction and imputation of the locus of control variable by using a list of alternative measures in Section 4.6 in which we i.a. also discuss potential attenuation bias due to state-dependent measurement error (see e.g Preuss and Hennecke, 2018).

4.3.3 Estimation Strategy

Using this data, we start the empirical analysis by estimating the effects of an individual's locus of control on her probability of being in a higher category of alcohol consumption as measured by the ordinal drinking variable in an ordered response model. In line with Long and Freese (2014), the ordered choice situation can be presented in form of a latent variable model with D_{it}^* being the latent continuous drinking variable of individual i at time t, which is divided into $J = 4$ ordinal categories such that

$$D_{it} = m \quad \text{if } \alpha_{m-1} < D_{it}^* \leq \alpha_m \quad \text{for } m = 1 \dots 4 \tag{4.1}$$

Where α_{m-1} and α_m are the lower and upper cut-points with $\alpha_0 = -\infty$ and $\alpha_m = \infty$. The probability that a drinking outcome m occurs is given by

$$Pr(D_{it} = m|\boldsymbol{X_{it}}) = Pr(\alpha_{m-1} < D_{it}^* \leq \alpha_m|\boldsymbol{X_{it}}) \text{ for } m = 1 \dots 4 \tag{4.2}$$

with $\boldsymbol{X_{it}}$ being the vector of regressors. The vector includes the (imputed) locus of control loc_{it} of individual i in t. In order to identify potential non-linearities in the effects, in addition to the continuous variable loc_{it}, the LOC is alternatively also included as a categorical variable that indicates in which tercile of the LOC distribution an individual is classified. Additionally, $\boldsymbol{X_{it}}$ contains an extensive list of demographic information C_{it} (gender, age, nationality, region of residence, number of children in the household, family status), educational controls (school degree, vocational degree and college degree), controls for the economic situation (net household income, gross labor income, labor force status and occupational autonomy) and health-related controls (disability-status, subjective health and body mass index). Furthermore, averaged and standardized personality and preferences measures P_i are included, which are the Big Five personality traits, general and health-related risk aversion as well as patience as a proxy for individual time-preferences, and T captures year fixed-effects. See Table 4.A.1 for the full list of controls.
Equation 4.2 is estimated using an ordered logit model, such that

$$\begin{aligned} Pr(D_{it} = m) &= Pr(\alpha_m < \boldsymbol{X_{it}}'\boldsymbol{\beta} \leq \alpha_{m-1}) && (4.3) \\ &= G(\alpha_m - \boldsymbol{X_{it}}'\boldsymbol{\beta}) - G(\alpha_{m-1} - \boldsymbol{X_{it}}'\boldsymbol{\beta}) && (4.4) \end{aligned}$$

with $\boldsymbol{X_{it}}'\boldsymbol{\beta} = \beta_1 loc_{it} + \beta_2 C_{it} + \beta_3 P_i + \beta_4 T$ and $G(\cdot)$ being the cumulative distribution function of a standard logistic distribution. Estimated coefficients are reported in log-odds units. Standard errors are clustered on the personal level in order to account for the serial correlation in the error terms due to the panel-nature of the data.

This ordered response model is based on the proportional odds or parallel regression assumption, which implies that the β-coefficient – and consequently also the effects of the explanatory variables on the drinking variable – are constant for all $J - 1$ binary decision situations. The slopes of all regressions are identical and the regression lines are parallel as only the intercepts differ across values of J. As Long and Freese (2014) note, this assumption is violated very frequently and the ordered response model is therefore often assumed to be overly restrictive. To check, whether the ordered response model is the appropriate model for the present analysis, we conduct the Brant test, a Wald test for the parallel regression assumption (Brant, 1990), as an omnibus test for the entire model as well as separately for each of the independent variables in the model.

Based on the results of the Brant test, we then investigate the non-parallel effects by splitting the ordinal variable into the four binary indicators as discussed in Section 4.3.1. We estimate the average marginal effects of an individual's locus of control on her probability of being (1) a moderate or heavy drinker as opposed to being an abstainer or rare drinker (D_{1it}), (2) a moderate drinker as opposed to being an abstainer or rare drinker (D_{2it}), (2) a heavy

Table 4.3: Descriptive Analysis - Mean Comparison

	All		**Women**		**Men**	
	Ext.	Int.	Ext.	Int.	Ext.	Int.
	(1)	(2)	(3)	(4)	(5)	(6)
Alcohol Consumption						
Abstainers	0.15	0.09***	0.18	0.12***	0.11	0.06***
Rare Drinkers	0.30	0.27***	0.37	0.34***	0.22	0.19***
Moderate Drinkers	0.39	0.45***	0.37	0.43***	0.42	0.46***
Heavy Drinkers	0.16	0.19***	0.08	0.10***	0.25	0.28***
Moderate or Heavy Drinking D_1 (Binary Indicators)						
All	0.55	0.64***	0.45	0.54***	0.67	0.74***
Beer	0.36	0.42***	0.18	0.21***	0.57	0.62***
Wine	0.34	0.43***	0.37	0.46***	0.32	0.41***
Spirits	0.10	0.11***	0.06	0.07	0.15	0.15
Mixed Drinks	0.05	0.06**	0.05	0.07***	0.06	0.06
Observations	17,903	18,231	9,651	9,065	8,252	9,166

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Individuals are grouped into internals (Int.) and externals (Ext.) based on whether their LOC is lower/equal or higher than the median. Significance stars refer to significance level of t-test for mean equivalence between externals and internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

drinker as opposed to being a moderate drinker (D_{3it}) and (4) a heavy drinker as opposed to being an abstainer or rare drinker (D_{4it}):

$$P(D_{it} = 1) = P(\beta_1 + \beta_2 loc_{it} + \beta_3 C_{it} + \beta_3 P_i + \beta_4 T + \epsilon_{it} > 0), \tag{4.5}$$

where D_{it} is either one of the four indicators for alcohol consumption of individual i at time t. Equation 4.5 is estimated using a simple logit approach. Standard errors are again clustered on the personal level and the results presented in Section 4.4 are the average marginal effects. The locus of control is always standardized and categorized within the sub-samples such that e.g. having a high LOC corresponds to a high LOC as compared to all other individuals within the two selected consumption categories and within the same gender.

4.4 Results

4.4.1 Descriptive Evidence

Table 4.3 summarizes the results of a first descriptive analysis of the relationship between LOC and alcohol consumption. The share of individuals who indicate a certain frequency of alcohol consumption are computed separately for internals (Int.) and externals (Ext.) with the cutoff being the median for this simple aggregation. The results of the t-tests for mean equality indicate that for men as well as women the share of individuals who indicate that they are moderate or heavy drinkers is significantly higher in the group of internal individuals. While for the full sample of internals the share of moderate drinkers is 45%, only 39% of externals are

moderate drinkers. In this raw difference, internals are also more likely to be heavy drinkers (19% as opposed to 16%). All differences similarly hold for men and women. In line with this, internals are also less likely to be abstainers (9% as opposed to 15%). The significant difference applies to all types of beverage in the full sample and is thus not driven by a certain type of alcohol. Nevertheless, this descriptive relationship is very likely to be driven by a long list of socio-economic characteristics that are associated with a higher consumption probability and a higher LOC, such as education, age and earnings.

4.4.2 Main Results

Based on this first descriptive evidence, Table 4.4 summarizes the estimated coefficients of the ordered response model using the categorical drinking variable as the dependent variable.

The descriptive raw gap between internals and externals remains significantly positive also if additional sets of control variables are added to the model. Nevertheless, including those controls, one set at a time, reveals that the raw gap was especially driven by educational and economic variables being omitted (columns (1) and (2) compared to columns (3) and (4)) while e.g. including personality and preferences measures does not strongly effect the LOC estimates (columns (5) and (6) compared to columns (7) and (8)). Columns (7) and (8) contain the results for the full specification and are referred to in the following.[7] An increase in the individual's LOC by one standard deviation on average increases the log odds of being in a higher level of alcohol consumption by 0.07 holding all other variables constant. If we instead consider the tercile dummies as explanatory variables, no severe non-linearity in the effects can be identified. While having a medium LOC ($(LOC_{P33}, LOC_{P66}]$) on average increases the log odds by 0.13 as compared to having a low LOC ($[LOC_{min}, LOC_{P33}]$), having a high LOC ($(LOC_{P66}, LOC_{max}]$) increases the log odds by 0.162. Hence, locus of control is associated with a higher probability of being in a higher category of alcohol consumption, i.e. higher drinking frequencies, also if we control for a large set of potentially important endogenous factors such as education, health or other personality traits.

Nevertheless, the size of the effect cannot be given a reasonable interpretation. Additionally, the test statistics of the Brant test for parallel regressions indicates a strong violation of the proportional odds assumption in the present case. The test statistics of the test for the full model are significant throughout all specifications (see bottom of each panel in Table 4.4). Additionally, Table 4.A.2 reports the test statistics for each regressor separately in the example of the full specification (column (7) in Table 4.4). The test statistics are also highly significant for a large number of control variables such as e.g. gender, age, nationality, vocational degree, gross labor income, health, BMI, agreeableness, risk aversion and patience but most importantly locus of control. This indicates that the hypothesis of equal β-coefficients can be

[7] The full estimation results for the specifications in column (7) can be found in Table 4.A.2 in the Appendix.

Table 4.4: Main Results (Ordered Logit Coefficients) - Categorical Drinking Variable

	Outcome Variable: Categorical Drinking Variable							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All								
LOC Factor (cont.)	0.173***		0.090***		0.059***		0.070***	
	(0.014)		(0.014)		(0.014)		(0.015)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.289***		0.170***		0.126***		0.130***
		(0.030)		(0.030)		(0.030)		(0.031)
$(LOC_{P66}, LOC_{max}]$		0.383***		0.201***		0.137***		0.162***
		(0.032)		(0.032)		(0.032)		(0.034)
Observations	36,134	36,134	36,134	36,134	36,134	36,134	36,134	36,134
Brant Test								
chi2	1076.806	1047.331	1410.956	1398.826	1431.146	1423.392	1546.511	1540.867
Brant $p > chi2$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Men								
LOC Factor (cont.)	0.129***		0.065***		0.041**		0.056***	
	(0.020)		(0.020)		(0.020)		(0.021)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.250***		0.148***		0.116***		0.126***
		(0.044)		(0.044)		(0.044)		(0.045)
$(LOC_{P66}, LOC_{max}]$		0.289***		0.148***		0.103**		0.135***
		(0.045)		(0.046)		(0.046)		(0.049)
Observations	17,418	17,418	17,418	17,418	17,418	17,418	17,418	17418
Brant Test								
chi2	334.568	316.539	533.773	525.065	555.237	548.821	623.704	618.986
$p > chi2$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Women								
LOC Factor (cont.)	0.210***		0.107***		0.069***		0.074***	
	(0.019)		(0.019)		(0.019)		(0.021)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.279***		0.147***		0.094**		0.088**
		(0.041)		(0.042)		(0.042)		(0.042)
$(LOC_{P66}, LOC_{max}]$		0.457***		0.226***		0.144***		0.155***
		(0.044)		(0.045)		(0.045)		(0.048)
Observations	18,716	18,716	18,716	18,716	18,716	18,716	18,716	18,716
Brant Test								
chi2	541.260	534.497	758.872	759.394	759.394	761.473	777.585	778.618
$p > chi2$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Year Fixed-Effects	✓	✓	✓	✓	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓	✓	✓	✓	✓
Educational Controls			✓	✓	✓	✓	✓	✓
Economic Controls			✓	✓	✓	✓	✓	✓
Health Controls					✓	✓	✓	✓
Personality Controls							✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Full estimation results for column (7) as well as Brant test results for each of the independent variables in the model can be found in Table 4.A.2 in the Appendix. All other full results as well as Brant test statistics are available from the authors upon request.

Table 4.5: Main Results (Logit, Marginal Effects) - Binary Drinking Indicators

	Full Sample		**Selected Samples**					
	Moderate/Heavy Drinking D_1		Moderate Drinking D_2		Heavy Drinking D_3 *(vs. Moderate)*		Heavy Drinking D_4 *(vs. Non/Rare)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All								
LOC Factor (cont.)	0.015***		0.017***		0.003		0.009**	
	(0.003)		(0.004)		(0.004)		(0.004)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.025***		0.024***		-0.006		0.022**
		(0.007)		(0.008)		(0.009)		(0.009)
$(LOC_{P66}, LOC_{max}]$		0.036***		0.041***		-0.002		0.023**
		(0.008)		(0.009)		(0.010)		(0.010)
Observations	36,134	36,134	29,736	29,736	21,574	21,574	20,958	20,958
Men								
LOC Factor (cont.)	0.014***		0.018***		0.006		0.009	
	(0.005)		(0.006)		(0.006)		(0.007)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.028***		0.030**		-0.011		0.040***
		(0.010)		(0.012)		(0.013)		(0.015)
$(LOC_{P66}, LOC_{max}]$		0.031***		0.037***		-0.001		0.031*
		(0.011)		(0.014)		(0.014)		(0.017)
Observations	17,418	17,418	12,712	12,712	12,342	12,342	9,782	9,782
Women								
LOC Factor (cont.)	0.015***		0.016***		0.001		0.008*	
	(0.005)		(0.005)		(0.005)		(0.005)	
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)								
$(LOC_{P33}, LOC_{P66}]$		0.018*		0.018*		-0.003		0.008
		(0.010)		(0.011)		(0.011)		(0.010)
$(LOC_{P66}, LOC_{max}]$		0.037***		0.040***		-0.005		0.017
		(0.012)		(0.012)		(0.013)		(0.011)
Observations	18,716	18,716	17,024	17,024	9,232	9,232	11,176	11,176
All Controls	✓	✓	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Full estimation results for column (1) can be found in Table 4.A.3 in the Appendix. All other full estimation results are available from the authors upon request.

rejected for a large number of explanatory variables and the parallel regression assumption is strongly violated in the ordered response model.

In order to estimate non-parallel effects and additionally be able to give the effects a reasonable interpretation, Table 4.5 summarizes the estimated average marginal effects of locus of control using the binary indicators as the dependent variable. All estimations are using the full specification corresponding to columns (7) and (8) in Table 4.4.[8] Columns (1) and (2) include the estimated marginal effects on our main drinking indicator for moderate or

[8] The full estimation results for the specifications in columns (1), (3), (5) an (7) can be found in Table 4.A.3 in the Appendix.

heavy consumption, i.e. the indicator which uses the full sample. An increase in the individual's LOC by one standard deviation on average increases the probability of moderate or heavy consumption by 1.5 percentage points holding all other variables constant. While having a medium LOC ($(LOC_{P33}, LOC_{P66}]$) on average increases the probability of occasional or regular consumption by 2.5 percentage points as compared to having a low LOC ($[LOC_{min}, LOC_{P33}]$), having a high LOC ($(LOC_{P66}, LOC_{max}]$) increases the probability by 3.6 percentage points. This effect is of considerable economic relevance. Given the overall share of moderate and heavy drinkers of about 60 percent the effect of a high as opposed to a low LOC translates into a 6 percent increase in probabilities. As can be seen in Table 4.A.3, the effect size is comparable with the effects of other important variables such as number of children in the household, university degree or employment status and higher than the effect of e.g. household income, body mass index and, surprisingly, of similar size as the marginal effect of important preferences measures such as willingness to take risks (general and health-related) and patience as a proxy for time-preferences. No severe gender differences can be observed for the absolute effect of both the continuous as well as the dichotomous LOC measure. While the effect seems to be largely linear for women, the effect flattens out for men as the effect sizes for a medium and a high LOC are not significantly different from each other. Nevertheless, given the lower overall share of drinkers in the sample of women, the relative effects are larger for women (7.5% as opposed to 4.4% increase for a high LOC).

Additionally, columns (3) to (8) give the average marginal effects for the alternative indicators which only compare two of the three groups while setting the third to missing.[9] These indicators enable us to make statements about the effects on the probabilities of moderate and heavy consumption independent from each other. Nevertheless, it has to be noted that the estimated effects using those outcome variables are at risk of being biased by sample selection. The estimated results thus have to be interpreted with care and are only ancillary evidence in addition to the main indicator. Additionally, the locus of control is always standardized and categorized within the sub-samples such that e.g. having a high LOC corresponds to a high LOC as compared to all other individuals within these two selected consumption categories. If we abstract from potential sample selection bias, comparing the results reveals that an internal LOC increases the probability of being a moderate drinker (columns (3) and (4)) as well as also the probability of being a heavy drinker (columns (7) and (8)) as compared to being an abstainer or rare drinker for the sample of all individuals. The effect is considerable stronger for moderate drinking, which is nevertheless mainly driven by the lower share of heavy drinkers in the population. Having a high LOC on average increases the probability of being a moderate drinker by 4.1 percentage points (approx. 10%) whereas the marginal effect on heavy drinking is 2.3 percentage points (approx. 13%). If we instead compare heavy drinking to moderate drinking, LOC cannot be found to have a significant effect on being a

[9] Abstainers and rare drinkers are always grouped into one reference group. An analyses which compares rare drinkers to abstainers is presented in the robustness checks in Section 4.6.2.

heavy drinker (columns (5) and (6)). An important heterogeneity can be found with respect to gender. While LOC increases the probability of moderated drinking similarly for men and women, the effect on heavy drinking is largely driven by men while the effect is distinctly lower and insignificant for women if the dichotomous LOC indicators are used. Additionally, the effect for men seems to be largely nonlinear with the largest effect being observable for a medium LOC.

4.5 Theoretical Considerations

In line with first considerable doubts about the applicability of the health investment model onto the relationship between locus of control and alcohol consumption raised by Cobb-Clark *et al.* (2014) which is based on their findings of a positive effect of LOC on binge drinking, our empirical results stand in stark contrast to the existing findings on the effect of locus of control on (risky) health behavior in other domains such as smoking, exercise and healthy diets. Thus, the considerations by Cobb-Clark *et al.* (2014) on the missing subjective link between current alcohol consumption and future health consequences can be supported. In line with Bennett *et al.* (1998), alcohol consumption might be associated with higher amounts of uncertainty of future outcomes as individuals might not see alcohol consumption in moderate amounts as affecting their health too strongly. This is also in accordance with the discourse about positive and negative health effects of alcohol consumption in the medical literature which was summarized in Section 4.1. Although individual considerations about health investments might nevertheless be at play, they are on average dominated by other mechanisms in the analysed population. It is thus the goal of this section to discuss two likely mechanisms which might explain the estimated relationship. Hence, in line with the existing literature discussed in Section 4.2, locus of control might positively affect alcohol consumption (1) through the increased considerations about the investment in social and occupational networks and the peer effects of alcohol consumption as well as (2) through the decreased perception of risks driven by information constraints and cognitive limitations.

4.5.1 Investment into Social and Occupational Networks

Based on the existing psychological literature on peer effects of alcohol consumption especially in adolescence (Duncan *et al.*, 2005; Argys and Rees, 2008; Lundborg, 2006), a first likely mechanism might be the link via differences in the importance and in the existence of peer and networking effects. Alcohol consumption might be associated with important positive effects on social and occupational networks as drinking is often inevitably connected to social gatherings and abstinence has been shown to be linked to strong negative penalties with respect to social integration (see e.g. Leifman *et al.*, 1995; Buonanno and Vanin, 2013). For example, Peters and Stringham (2006) as well as Ziebarth and Grabka (2009) discuss the association between alcohol consumption and social networks as likely channels for their

Table 4.6: Additional Results (Marginal Effects) - Social Network (Outcome: Moderate or Heavy Drinker)

	All		**Men**		**Women**	
	(1)	(2)	(3)	(4)	(5)	(6)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)						
$(LOC_{P33}, LOC_{P66}]$	0.024***	0.015**	0.025**	0.018*	0.016	0.006
	(0.008)	(0.007)	(0.010)	(0.010)	(0.011)	(0.010)
$(LOC_{P66}, LOC_{max}]$	0.038***	0.029***	0.031***	0.022*	0.039***	0.029**
	(0.008)	(0.008)	(0.011)	(0.011)	(0.012)	(0.012)
Spare Time Activities (Imputed, Categorical)						
Going out Eating/Drinking		0.050***		0.040***		0.051***
		(0.004)		(0.006)		(0.006)
Visiting Neighbors and Friends		0.003		0.003		0.008
		(0.004)		(0.006)		(0.006)
Attend Social Gatherings		0.039***		0.040***		0.039***
		(0.004)		(0.005)		(0.006)
Attend Cultural Events		0.026***		0.036***		0.037***
		(0.005)		(0.007)		(0.007)
Observations	35,246	35,246	16,992	16,992	18,254	18,254
All Controls	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

identified positive effect of alcohol consumption on earnings. As they notice, alcohol consumption still is a social norm in modern Western societies which inevitably links drinking and the attendance of social gatherings. Thus, moderate drinking produces social capital and can be labeled a productive activity. In line with the argumentation about locus of control and investment into future outcomes (see e.g. Caliendo *et al.*, 2015b; Coleman and DeLeire, 2003; Cobb-Clark *et al.*, 2014), internals are expected to invest more into social capital and networks by attending social gatherings and, in doing so, drinking. This is why they might, by default, be more likely to drink alcohol in moderation.

The estimates, which are presented in Table 4.6, are intended to investigate whether internals are simply more likely to be exposed to alcohol during social gatherings. This is done by controlling for different categories of spare time activities such as going out eating and drinking, attending social gatherings, visiting neighbours and friends and attending cultural events as proxies for investments into social networks. Individuals rate the frequency with which they participate in those activities on a scale from 1 (Weekly) to 4 (Never). The scale is reversed in order to indicate a frequent participation for high values and included as a continuous variable into the estimations for reasons of simplification and in order to prevent a loss of information.[10] As the activities are surveyed irregularly, they are imputed into the

[10] In the years 2008 and 2013 the scale included a category "daily". This is considered as an additional value 5 for those years.

relevant years and are thus only proxies for the actual level of social activity in the survey years.[11] Table 4.6 gives an overview over the estimated results of this additional analysis. The process of imputation is accompanied by a loss in observations. In order to separate the changes in effect size due to changes in the sample and inclusion of controls, the baseline results where replicated using the reduced sample. The results are not substantially different and are presented in columns (1), (3) and (5) of Table 4.6.

The estimation results of the additional analysis presented in columns (2), (4) and (6) of Table 4.6 indicate that for men as well as women large parts of the estimated effects are indeed driven by different levels of participation in events and social activities. The overall effect drops from 2.4 (3.8) to 1.5 (2.9) percentage points for a medium (high) LOC. Thus, around 37.5 (23.7) percent of the overall effect are driven by different levels of social activity between internals and externals.[12] This share is relatively similar for men and women. Nevertheless, the effect remains significantly positive for the indicator of a high LOC for both sexes even if the different levels of social activity are controlled for. Different levels of social interactions in the individuals spare time are thus likely to be not the only driving factor behind the estimated differences.

4.5.2 Underestimation of Risks

As discussed above, the uncertainty about the direct link between the behavior and negative future outcomes on health is likely to be higher for alcohol consumption than for other domains of risky health behavior such as smoking, unhealthy diets and a lack of exercise. In line with the seminal work of Von Neumann and Morgenstern (1944) and Kahneman and Tversky (1979) on the effect of uncertainty on decision making, individual perceptions are highly important in those situations. Individuals have to build own expectations about the probabilities with which their behavior is associated with certain outcomes. In the present case, individuals have to estimate the likelihood with which their level of consumption entails negative future consequences for their health. The accuracy of these estimations is importantly affected by individuals cognitive resources and the information available to them (see e.g. Kenkel, 1991; Farrell and Fuchs, 1982; Cutler and Lleras-Muney, 2010) as well as their perception about their own susceptibility to health problems (see e.g. Weinstein, 1984, 1987; Janz and Becker, 1984). Thus, the empirical literature relatively consistently finds that individuals on average tend to be overly optimistic especially about their own risks and underestimate the negative consequences of e.g, smoking (Viscusi, 1990).

[11] "Going out eating and drinking" and "visiting neighbors and friends" (observed in 2008) is imputed forward and backward into all survey years 2006, 2008 and 2010 from the observation in 2008, "attending social gatherings" (observed in 2005, 2007 and 2009) is imputed forward from 2005 into 2006, from 2007 into 2008 and from 2009 into 2010 and "attending cultural events" (observed in 2005, 2007, 2008 and 2009) is imputed forward from 2005 into 2006 and from 2009 into 2010.

[12] It has to be noted that the results have to be interpreted with care as the inclusion of these endogenous controls in the analysis is likely to introduce selectivity bias (see e.g. Angrist and Pischke, 2008, p.66).

Various studies in medicine and psychology have already found that also locus of control has an important effect on individual perceptions about the personal risk e.g. with respect to myocardial infarction, AIDS, cancer and mortality (see e.g. Stürmer *et al.*, 2006; Rosengren *et al.*, 2004; Bosma *et al.*, 1999; Hoorens and Buunk, 1993; Cull *et al.*, 1999; Frijling *et al.*, 2004; Källmén, 2000; Sjoberg, 2000). For example Hoorens and Buunk (1993) show that students with an internal LOC are more likely to report a lower personal risk of having drinking problems. Additionally, Becker *et al.* (2012) find a significant positive association between LOC and subjective health. In line with this literature also Cobb-Clark *et al.* (2014) argue that an increased perception of control might be correlated with a higher belief about the ability to cope with and prevent the consequences of drinking. An increased perception of individual control (being internal) might thus decrease the perceived importance of risk for life's outcomes. The future risks of alcohol consumption might thus be underestimated if the individual control is overestimated. The relationship between LOC and risk perceptions has first been discussed in Slovic (1992).

This argument is largely in line with the latest literature on the association between an internal LOC and risky and inconsistent investment decisions (Salamanca *et al.*, 2016; Pinger *et al.*, 2018). Salamanca *et al.* (2016) find that household heads with an internal LOC are more likely to invest in equity. They assume that this is driven by a lower perceived variance in equity and thus a lower perceived risk of these investments for internal individuals. Additionally, Pinger *et al.* (2018) find that internals are also more likely to make inconsistent investment decisions driven by a higher probability of corresponding to a hot hand fallacy.

In line with Salamanca *et al.* (2016) we can show that LOC is associated with risky behavior over and beyond risk and time preferences. This is important to note as LOC has already been shown to be strongly associated with time preferences (Plunkett and Buehner, 2007; Becker *et al.*, 2012) and in line with the findings by Becker *et al.* (2012) the results presented in Table 4.B.1 in the Supplementary Material indicate the same important relationship also for risk aversion[13].

In order to assess the reasonableness of the supposed relationship between LOC, risk perception and alcohol consumption, the heterogeneity of the estimated effects with respect to the underlying risk aversion is empirically analysed. The reasoning behind this analysis is the assumption that the effect of differences in perceived risks is likely to be more pronounced for individuals with a higher risk aversion. If an individual has a high willingness to take risks, higher perceived risks might not change the decision making as strongly.

[13] While the general willingness to take risks is significantly higher for internals for both men and women, an interesting difference between general and health-related willingness to take risks can be seen. While the willingness to take health risks is also significantly higher for internal men, it is significantly lower for internal women. Thus, while for men general and health related risk taking don't seem to differ, LOC has an adverse effect on general and health-related risk aversion for women. Additionally, the difference between general and health-related risk attitude is significantly higher for internals in both genders. Thus, internals can be assumed to be more health-conscious than externals if the willingness to take health risks is weighted by the overall higher willingness to take risk.

Table 4.7: Additional Results (Marginal Effects) - Heterogeneity Analysis by Risk-Aversion (Outcome: Moderate or Heavy Drinker)

	All	**Men**	**Women**
	(1)	(2)	(3)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$ × Low Risk Taking	0.028***	0.036**	0.020
	(0.010)	(0.015)	(0.013)
× High Risk Taking	0.023**	0.020	0.013
	(0.011)	(0.013)	(0.018)
$(LOC_{P66}, LOC_{max}]$ × Low Risk Taking	0.049***	0.045***	0.042***
	(0.011)	(0.017)	(0.014)
× High Risk Taking	0.023**	0.019	0.027
	(0.012)	(0.014)	(0.019)
Observations	36,134	17,418	18,716
All Controls	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Thus, Table 4.7 gives an overview over the results of this heterogeneity analysis. The reported marginal effects are based on a model in which LOC is interacted with a binary indicator for high willingness to take risk ("high risk taking") based on the median as a cutoff and are computed separately for individuals with a low and high willingness.

It is obvious from the estimation results that the effect of LOC on alcohol consumption is highly heterogeneous with respect to underlying willingness to take risk. The positive effect of a medium and high LOC is considerably stronger for individuals with a low general willingness to take risk for both males and females. This is in line with the expectations and thus further supports the idea about risk perception being an important channel.

4.5.3 Subgroup Analysis - Occupation and Educational Level

In order to draw a clearer picture about the effect of LOC on alcohol consumption and the population groups for which we observe it, we conducted a subgroup analysis. Identifying heterogeneity between population groups might assist us in identifying the likely mechanism behind the estimated effects. Table 4.8 gives an overview over the marginal effects estimated separately for combinations of educational level and occupational autonomy. Individuals are grouped into the nine groups based on whether they have 1) a low education (less than 12 years), 2) medium education (12 - 14.5 years) or 3) high education (15 or more years) and whether they have 1) no or a low occupational autonomy (e.g. unemployed individuals or unqualified workers), (2) a medium occupational autonomy (e.g. overseers or qualified occupations) or 3) a high occupational autonomy (e.g. foremen and managers). Educational level and occupational autonomy are combined in order to potentially distinguish between occupational effects and cognitive abilities. All estimations include the full specification using

Table 4.8: Additional Results (Marginal Effects) - Subgroup Analysis (Outcome: Moderate or Heavy Drinker)

	Low Education			**Medium Education**			**High Education**		
	Occupational Autonomy			Occupational Autonomy			Occupational Autonomy		
	Low	Med	High	Low	Med	High	Low	Med	High
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)									
$(LOC_{P33}, LOC_{P66}]$	0.022*	0.014	0.020	0.030	0.004	-0.032	-0.015	-0.051	0.033*
	(0.012)	(0.022)	(0.037)	(0.021)	(0.026)	(0.033)	(0.029)	(0.032)	(0.020)
$(LOC_{P66}, LOC_{max}]$	0.057***	0.042*	0.044	0.009	0.008	0.018	-0.004	-0.005	0.011
	(0.013)	(0.024)	(0.038)	(0.022)	(0.029)	(0.037)	(0.031)	(0.035)	(0.022)
Observations	14,866	3,679	1,275	4,495	2,563	1,411	2,132	1,616	4,097
All Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the moderate or heavy consumption indicator and the discrete LOC. In order to take care of potential selection, the LOC is recalculated within each subgroup such that e.g. a high LOC always indicates a LOC higher than the 66th percentile within the group.

Due to in parts very low sample sizes, significance levels get lost for most subgroups which only allows for a very careful interpretation of the resulting effects. The results nevertheless indicate important heterogeneity. Overall, the effects are distinctly larger for individuals with a low education pointing into the direction that cognitive limitations and information constraints while processing the potential risks of alcohol consumption might actually play an important role. For individuals with the maximum level of cognitive ability (as measured by education) the effects are remarkably lower and again highly heterogeneous. Highly educated individuals with a high occupational autonomy seem to be the only ones who also show an increased probability to consume alcohol through an internal LOC. This might again point into the direction of investments into occupational networks through the attendance of events at which alcohol is consumed. The internal manager who is interested in expanding her occupational networks might thus drink one or the other glass of wine more on all the events she attends. Also the subgroup analysis thus indicates an important role of both proposed channels.

4.6 Robustness Checks

In addition to the ancillary estimations in Section 4.5, which had the goal of drawing a clearer picture of the mechanisms behind the identified effects, the sensitivity of the estimated effects with respect to a number of measurement issues in the dependent as well as explanatory variable and potential confounders has to be discussed and analysed in detail in the following section.

4.6.1 Locus of Control

In order to test the robustness of the estimated effects with respect to the construction and imputation of the locus of control measure, a number of alternative LOC measures are generated and the estimated effects are re-estimated for these alternative explanatory variables. Table 4.A.4 in the Appendix contains the results for these sensitivity checks. All models are estimated using the full specification and the indicator for moderate or heavy drinking as the outcome variable. The baseline results in the top panel of Table 4.A.4 thus refer to the results in column (2) of Table 4.5. Instead of using the weights from the confirmatory factor analysis, the simple LOC index, used in the estimations presented in panel (1), assumes equal weights of all 8 items used thus a simple average over all item responses is calculated. The results differ only marginally from those in the baseline. The only considerable changes is observed for the effect of a medium LOC for men. The robustness check presented in Panel (2) corresponds to the potential problem with inaccurate and volatile measurement of LOC in the surveys motivated by the discussion in e.g. Preuss and Hennecke (2018). In line with their findings on temporary effects of unemployment on LOC, the self-reported LOC is likely to depend on a lot of major and minor environmental and situational factors. If these state-effects are assumed to have no behavioral consequences and are pure noise, they are likely to lead to attenuation bias in the estimated effects when LOC is observed in different states across the sample. In order to investigate this problem, the explanatory variable is adjusted to an averaged LOC factor instead of the forward imputed version in the baseline in panel (2). The LOC factor is calculated as usual for each year but is imputed into all years using the average over the LOC observations from 2005 and 2010 for an individual. This adjustment is expected to reduce measurement error in the situational measurement of LOC. As expected, the estimated effects are considerably stronger especially for a high LOC, indicating a downward bias, likely driven by attenuation, in the baseline results. As has already been discussed in Section 4.3, concerns about reverse causality are rather minor in the case of alcohol consumption. Nevertheless, it might be interesting to see whether the effects are robust against the use of a LOC measure with lies even further in the past. Such a measure on the one hand might reduce reverse causality concerns but on the other hand again increases a potential attenuation bias as it is only a vague proxy for the LOC at time t. Panel (3) presents the estimation results of a model in which the LOC from the year 2005 is imputed forward into all observation years. The estimated effects do decrease especially for men but are still significant and of considerable size for the full sample.

4.6.2 Outcome Variable

Alternative Indicators In order to test the robustness of the estimated effects with respect to the choice of the drinking indicators as the main outcome variables, a number of alternative dichotomous measures of consumption are constructed and the estimated effects are

re-estimated for these alternative dependent variables. Table 4.A.5 in the Appendix contains the results for these sensitivity checks. All models are estimated using the full specification and the dichotomous LOC measure. The baseline results in the top panel of Table 4.A.5 thus refer to the results in column (2) of Table 4.5. In order to investigate whether effects are driven by differences at the extensive or intensive margin, they are re-estimated using an indicator for being an abstainer in panel (1a) as abstainers and rare drinkers where pooled in all main drinking indicators. In line with the main results, the effect is significantly negative, indicating a decreased probability of being an abstainer for internals. Additionally, panel (1b) adds to this by using an indicator for rare consumption as opposed to abstaining (with moderate and heavy consumption being set to missing). In line with expectations abstainers and rare drinker do not significantly differ with respect to their LOC, which is why pooling them into one reference group was suitable.

Objective Amounts More importantly, the estimated results might be biased by the subjective and rather vague nature of the alcohol consumption level available in the years 2006, 2008 and 2010. As our measure for alcohol consumption in the empirical analysis is based on the self-assessed amount of consumption, rated on a scale from "never" to "regular", the answer to this question not only depends on the actual consumption level but also on the individual perception of the terms regular and occasional. If individuals perceive amounts differently based on their locus of control, this would bias our results. As we know from the existing literature that internals are more health-conscious, for them only a few drinks might already be too much. They might thus only report higher levels of alcohol consumption without actually consuming more. The fact that we are able to observe significant effects also on abstaining from alcohol already is a first indication that differences in the reporting are not solely responsible for our result if we assume honestly reporting "never" to be largely objective. In order to further eliminate the bias of the estimated effects by differences in the perception of amounts between internals and externals, the estimations are repeated using a more objective measure of alcohol consumption which is available in the SOEP in the year 2016. In 2016, individuals do not self-assess their consumption levels but report objective amounts and frequencies. Individuals are asked to rate how often they drink alcohol on a scale from 1 (every day) to 6 (never) and how many drinks they do consume in a day when they drink on a scale from 1 (One to two drinks) to 5 (Ten or more drinks). An overview over the descriptive statistics for these variables can be found in Table 4.B.3 in the Supplementary Material. Thus, the new dependent variables are generated based on the reported frequency of consumption and the reported consumption amount per consumption day. Locus of control is imputed from the wave 2015. The measurement of LOC in 2015 as well as the construction of the factor is equivalent to those in 2005 and 2010 which was described in detail in Section 4.3.

The results of this sensitivity check are reported in panel (3) of Table 4.A.5. In panel (3a) the two available information are combined. The binary indicator for drinking thus is one if the

individual reports more then two drinking episodes per week ("high frequency") or more than four drinks per episode ("high amount"). The sensitivity check indicates that the results from the baseline are surprisingly robust with respect to the type of reporting. As opposed to the expectations, the baseline results seem to be biased downwards by the subjective reporting. When the drinking indicator is separated into high amounts and high frequencies, it becomes clear that the estimated effects are purely driven by differences in the consumption frequencies (panel (3c)). This strongly supports the idea of differences in networking behavior and thus the frequency of drinking opportunities being an important driver behind the differences. Nevertheless, as the estimated effects in panel (3b) show, LOC has a significant effect also on consumption amounts for men. Hence, important gender differences in the underlying mechanisms might be at play.

4.6.3 Compensating Health Behavior

Finally, individuals might compensate health behavior in certain domains with behavior in other domains. Thus, if an individual already invests a lot of time and effort e.g. into doing sports and eating healthy, the individual might be more willing to drink one or the other drink more as she considers already doing "enough" for her health. Vice versa an individual who already has a relatively unhealthy lifestyle might reduce the amount of alcohol consumption as the maximum level of "unhealthyness" is already reached. If locus of control now is associated with an increased amount of investment into health in the domains diet, exercise and smoking, this might lead to more alcohol consumption for those individuals.

In order to assess the importance of compensating, the reported behavior in the other domains (smoking, diet and exercise) is included as controls into the estimation model. The health behavior in the other domains is included as ordinal variables with the value manifestations 0 for "non" (no smoking, no exercise, no health diet), 1 for "moderate/light" and 2 for "strong/heavy". An overview over the descriptive statistics for these other health behaviors can be found in Table 4.B.2 in the Supplementary Material. In line with the findings by Cobb-Clark *et al.* (2014), internal individuals are significantly less likely to obtain risky health behavior throughout all domains. The results of the sensitivity check are presented in Table 4.A.6 in the Appendix. Controlling for the behavior in the other health domains does hardly change the estimated effects of LOC on alcohol consumption, which indicates no omitted variable bias in this context.

4.7 Conclusion

Based on existing knowledge about the role of locus of control for the individual investment into future outcomes, a spontaneous ex ante prediction of the potential effect of LOC on levels of alcohol consumption pointed into the direction of an expected negative relationship.

Internal individuals should have higher subjective expectations about the effect of their own efforts into drinking less or abstaining on their future health. But they do not!

As opposed to this prediction, we find a significant positive effect of locus of control on alcohol consumption. Internal individuals are on average 10% (4.1 p.p.) more likely to be moderate drinkers and 13% (2.3 p.p) more likely to be heavy drinkers as opposed to external individuals. These findings are robust against controlling for an extensive list of control variables, the variation in the LOC construct as well as the outcome variable – including the use of a more objective measure of consumption – and also hold if the health behavior in other domains is included into the model as a potential confounder.

As has already been noted by Cobb-Clark *et al.* (2014), this obvious dissent with the existing literature on locus of control and health investments is likely to be driven by the fact that the link between behavior and potential negative future outcomes is more subject to uncertainty than other domains of health behavior. Alcohol might not be perceived as a behavior which strongly affects health outcomes in general. Thus, locus of control does not affect health investments in this domain. As opposed to this, other mechanisms seem to be at play which cause a positive relationship. Based on the earlier literature on the behavioral consequences of LOC as well as the existing knowledge on the drivers of drinking behavior, we discuss two potential channels. Firstly, large parts of the relationship might be driven by differences in the social behavior and investments into social networks between internal and external individuals. Individuals with an internal LOC invest into social and occupational networks more strongly by attending different social events. As drinking is often inevitably connected to social gatherings and abstinence has been shown to be linked to strong negative penalties with respect to social integration, internals might thus be more exposed to alcohol in general. An ancillary analysis, in which the frequency of a number of spare time activities which are associated to social networks are included as controls, indicate that indeed parts of the effect can be explained by different levels of social interaction. As a second potential channel, we discuss differences in the perception of risks between internals and externals. Based on the literature on financial investments, we suggest that internal individuals more strongly believe in or overestimate their ability to cope with and prevent the negative consequences of drinking. Thus, they might underestimate the risk which is associated with drinking. A heterogeneity analysis with respect to underlying willingness to take risks as well as a subgroup analysis, which shows important heterogeneity with respect to individual cognitive ability, supports these considerations.

We do find indication for both mechanisms. Moreover, differential effects of LOC on consumption frequencies and amounts, which are identified using an objective measure of alcohol consumption from a more recent SOEP wave, reveal that important gender differences in the underlying mechanisms might be at play. While the effect for women largely concerns consumption frequencies, men are also more likely to drink more if they are internal. Thus,

differences in the perception of risk might be especially important for men. Additionally, also considerations about health investments might still be at play in the alcohol domain, though they are overshadowed by the mechanisms discussed above. Further disentangling the effects with respect to the underlying channels is beyond the scope of this paper. Nevertheless, this might be an important path for future research as both mechanisms are expected to have distinct economic and medical consequences. Whereas drinking as an investment decision might improve the occupational and economic success while being related to rather moderate amounts of alcohol consumption, an underestimation of risks is potentially associated with regular and heavy episodic drinking and thus with the economic costs involved, e.g. through the strain it puts on individual health care expenditures and labor market perspectives.

Our paper significantly contributed to the literature on the explanation of health behavior and risky decision making. Given the consensus about the risks of alcohol consumption for individual health and economic outcomes, knowing about this specific intrinsic driver of drinking can e.g. crucially contribute to the targeting of interventions with the goal of reducing alcohol consumption in the population. Additionally, the paper also adds an interesting new light to the literature on the behavioral implications of the personality trait locus of control which was so far largely unified in the assumption that an internal locus of control is associated with mainly positive outcomes and desirable behaviors and decisions.

4.A Appendix

Table 4.A.1: Summary Statistics

	All (1)	**Men** (2)	**Women** (3)
Demographic Controls			
Female	0.52		
Age	46.64	46.92	46.38
German Nationality	0.94	0.94	0.94
East-Germany	0.27	0.27	0.27
Number of Children in HH	0.48	0.47	0.50
Married	0.65	0.65	0.64
Educational Controls			
Highest School Degree			
No School Degree	0.02	0.02	0.02
Lower Secondary School	0.30	0.32	0.28
Intermediary School	0.33	0.29	0.36
Highschool	0.30	0.32	0.28
Other School	0.06	0.06	0.06
Highest Vocational Degree			
No Vocational Diploma	0.17	0.15	0.20
Apprenticeship	0.45	0.47	0.43
Higher Technical College	0.25	0.24	0.26
College or University Degree	0.24	0.26	0.22
Labor Market Controls			
Net Household Income in KEUR	2.96	3.03	2.90
Gross Labor Income in KEUR	1.74	2.38	1.14
Occupational Autonomy			
Low	0.34	0.28	0.39
Medium	0.26	0.28	0.24
High	0.22	0.18	0.25
Labor Force Status			
Employed	0.06	0.06	0.06
Self-Employed	0.59	0.63	0.56
Unemployed	0.07	0.09	0.05
Not-Working	0.06	0.01	0.10
Pensioneer	0.18	0.18	0.17
In Education	0.02	0.02	0.02
Military Service	0.00	0.00	0.00
Maternity Leave	0.02	0.00	0.03
Health Controls			
Disabled	0.11	0.13	0.10
In Bad Health	0.16	0.15	0.17
Body Mass Index (imputed)	26.13	26.92	25.40
Personality Controls			
Conscientiousness (avg.)	5.88	5.81	5.94
Extraversion (avg.)	4.81	4.68	4.93
Agreeableness (avg.)	5.37	5.19	5.53
Neuroticism (avg.)	3.86	3.61	4.10
Openness (avg.)	4.49	4.40	4.58
Willingness to take general risk (general) (avg.)	4.51	4.95	4.10
Willingness to take health risk (avg.)	2.98	3.34	2.64
Observations	36134	17418	18716
Individuals	15482	7478	8004

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.

Table 4.A.2: Full Estimation Results and Brant Test Statistics (All Individuals)

		Brant Test	
	b/se	chi2	$p > chi2$
	(1)	(2)	(3)
LOC Factor (cont.)	0.070***	16.397	0.000
	(0.015)		
Year Dummies			
2008	0.015	18.103	0.000
	(0.016)		
2010	-0.066***	9.581	0.008
Demographic Controls			
Female	-0.978***	280.427	0.000
	(0.034)		
Age	0.024**	16.189	0.000
	(0.009)		
Squared Age	-0.000	8.789	0.012
	(0.000)		
German Nationality	0.809***	107.343	0.000
	(0.075)		
East-Germany	0.005	53.981	0.000
	(0.033)		
Number of Children	-0.110***	3.168	0.205
	(0.019)		
Married	-0.007	4.765	0.092
	(0.034)		
Educational Controls			
Highest School Degree (Ref: No Degree)			
Lower Secondary School	0.453***	2.879	0.237
	(0.124)		
Intermediary School	0.713***	7.559	0.023
	(0.124)		
Highschool	0.767***	10.138	0.006
	(0.126)		
Other School	0.406**	1.673	0.433
	(0.135)		
Highest Vocational Degree (Ref: No Diploma)			
Apprenticeship	0.132***	57.536	0.000
	(0.039)		
Higher Technical College	0.066	66.540	0.000
	(0.043)		
College or University Degree	0.222***	21.453	0.000
	(0.045)		
Economic Controls			
Net HH Income in KEUR	0.082***	7.469	0.024
	(0.012)		
Gross Labor Income in KEUR	-0.012	22.485	0.000
	(0.010)		
Labor Force Status (Ref: Unemployed)			
Employed	0.194**	1.145	0.564
	(0.074)		
Self-Employed	0.208*	1.572	0.456
	(0.087)		
Not-Working	-0.143	4.217	0.121
	(0.077)		
Pensioneer	-0.056	1.453	0.484
	(0.075)		
In Education	0.002	7.610	0.022
	(0.093)		
Military Service	0.289	2.354	0.308

Continued on next page...

... continued from previous page

		Brant Test	
	b/se	chi2	$p > chi2$
	(1)	(2)	(3)
	(0.220)		
Maternity Leave	-0.375***	1.588	0.452
	(0.093)		
Occupational Autonomy (Ref.: Low)			
Medium	0.017	3.284	0.194
	(0.052)		
High	0.018	7.326	0.026
	(0.045)		
Health Controls			
Disabled	-0.387***	11.908	0.003
	(0.049)		
In Bad Health	-0.368***	31.446	0.000
	(0.037)		
Body Mass Index	-0.040***	21.719	0.000
	(0.003)		
Personality Controls			
Conscientiousness (avg.)	-0.088***	2.045	0.360
	(0.018)		
Extraversion (avg.)	0.124***	0.335	0.846
	(0.017)		
Agreeableness (avg.)	-0.069***	47.522	0.000
	(0.018)		
Neuroticism (avg.)	0.009	12.963	0.002
	(0.018)		
Openness (avg.)	-0.014	1.653	0.438
	(0.018)		
Willingness to take general risk (avg.)	0.069**	30.652	0.000
	(0.022)		
Willingness to take health risk (avg.)	0.091***	4.738	0.094
	(0.019)		
Patience (avg.)	-0.058***	14.094	0.001
	(0.017)		
Observations	36,134		

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.A.3: Full Results (Marginal Effects) - Binary Drinking Indicators (All Individuals)

	Moderate/Heavy Drinking D_1	Moderate Drinking D_2	Heavy Drinking D_3 *(vs. Moderate)*	Heavy Drinking D_4 *(vs. Non/Rare)*
	(1)	(2)	(3)	(4)
LOC Factor (cont.)	0.015***	0.017***	0.003	0.009**
	(0.003)	(0.004)	(0.004)	(0.004)
Year Dummies				
2008	0.009**	0.006	-0.010*	0.012**
	(0.004)	(0.005)	(0.006)	(0.005)
2010	-0.007	-0.006	0.007	-0.012**
	(0.005)	(0.006)	(0.006)	(0.005)
Demographic Controls				
Female	-0.205***	-0.152***	0.202***	-0.305***
	(0.008)	(0.009)	(0.009)	(0.010)
Age	0.004**	0.002	-0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)
Squared Age	-0.000	0.000	0.000***	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
German Nationality	0.118***	0.111***	-0.048**	0.100***
	(0.016)	(0.016)	(0.020)	(0.016)
East-Germany	-0.009	-0.006	0.007	-0.016*
	(0.008)	(0.009)	(0.010)	(0.009)
Number of Children	-0.026***	-0.026***	0.008	-0.026***
	(0.004)	(0.005)	(0.006)	(0.005)
Married	-0.006	-0.005	0.004	0.002
	(0.008)	(0.009)	(0.010)	(0.009)
Educational Controls				
Highest School Degree (Ref: No Degree)				
Lower Secondary School	0.052**	0.041	-0.039	0.083**
	(0.025)	(0.028)	(0.038)	(0.033)
Intermediary School	0.098***	0.082***	-0.077**	0.142***
	(0.025)	(0.029)	(0.038)	(0.034)
Highschool	0.102***	0.077***	-0.103***	0.166***
	(0.026)	(0.030)	(0.040)	(0.037)
Other School	0.046*	0.036	-0.038	0.076*
	(0.026)	(0.031)	(0.043)	(0.039)
Highest Vocational Degree (Ref: No Diploma)				
Apprenticeship	0.031***	0.041***	0.031***	0.002
	(0.009)	(0.010)	(0.011)	(0.011)
Higher Technical College	0.008	0.016	0.031***	-0.012
	(0.010)	(0.011)	(0.012)	(0.012)
College or University Degree	0.044***	0.042***	-0.006	0.038***
Economic Controls				
	(0.011)	(0.012)	(0.012)	(0.013)
Net HH Income in KEUR	0.020***	0.016***	-0.010***	0.023***
	(0.003)	(0.003)	(0.003)	(0.003)
Gross Labor Income in KEUR	-0.002	-0.001	0.003	-0.004**
	(0.002)	(0.003)	(0.003)	(0.002)
Labor Force Status (Ref: Unemployed)				
Employed	0.039**	0.039**	-0.014	0.037**
	(0.017)	(0.020)	(0.020)	(0.019)
Self-Employed	0.032	0.027	-0.036	0.036
	(0.021)	(0.024)	(0.024)	(0.023)
Not-Working	-0.013	-0.014	-0.015	-0.007
	(0.018)	(0.019)	(0.024)	(0.021)
Pensioneer	-0.012	-0.015	-0.018	0.000
	(0.017)	(0.019)	(0.021)	(0.019)
In Education	0.002	0.018	0.047*	-0.033
	(0.023)	(0.025)	(0.027)	(0.026)
Military Service	0.087	0.108	0.013	0.079

Continued on next page...

... continued from previous page

	Moderate/Heavy Drinking D_1	Moderate Drinking D_2	Heavy Drinking D_3 *(vs. Moderate)*	Heavy Drinking D_4 *(vs. Non/Rare)*
	(1)	(2)	(3)	(4)
	(0.064)	(0.073)	(0.067)	(0.075)
Maternity Leave	-0.094***	-0.089***	0.076*	-0.121***
	(0.024)	(0.025)	(0.043)	(0.032)
Occupational Autonomy (Ref.: Low)				
Medium	0.004	0.012	0.018	-0.012
	(0.012)	(0.014)	(0.014)	(0.014)
High	0.007	0.014	0.016	-0.010
	(0.011)	(0.013)	(0.012)	(0.012)
Health Controls				
Disabled	-0.073***	-0.068***	0.029**	-0.071***
	(0.011)	(0.012)	(0.013)	(0.012)
In Bad Health	-0.084***	-0.091***	-0.010	-0.059***
	(0.008)	(0.009)	(0.011)	(0.009)
Body Mass Index	-0.009***	-0.008***	0.007***	-0.011***
	(0.001)	(0.001)	(0.001)	(0.001)
Personality Controls				
Conscientiousness (avg.)	-0.019***	-0.019***	0.012**	-0.018***
	(0.004)	(0.005)	(0.005)	(0.005)
Extraversion (avg.)	0.027***	0.024***	-0.018***	0.030***
	(0.004)	(0.004)	(0.005)	(0.005)
Agreeableness (avg.)	-0.003	0.005	0.029***	-0.020***
	(0.004)	(0.005)	(0.005)	(0.005)
Neuroticism (avg.)	0.003	0.004	0.007	-0.004
	(0.004)	(0.005)	(0.005)	(0.005)
Openness (avg.)	-0.003	-0.002	0.007	-0.008*
	(0.004)	(0.005)	(0.005)	(0.005)
Willingness to take general risk (avg.)	0.018***	0.022***	0.010*	0.004
	(0.005)	(0.005)	(0.006)	(0.006)
Willingness to take health risk (avg.)	0.017***	0.014***	-0.019***	0.025***
	(0.004)	(0.005)	(0.005)	(0.005)
Patience (avg.)	-0.012***	-0.009**	0.013***	-0.018***
	(0.004)	(0.004)	(0.005)	(0.004)
Observations	36,134	29,736	21,574	20,958

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.A.4: Robustness Checks (Marginal Effects) - Locus of Control (Outcome: Moderate or Heavy Drinker)

	All (1)	**Men** (2)	**Women** (3)
Baseline			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.026^{***}	0.028^{***}	0.018^{*}
	(0.007)	(0.010)	(0.010)
$(LOC_{P66}, LOC_{max}]$	0.036^{***}	0.031^{***}	0.037^{***}
	(0.008)	(0.011)	(0.012)
Observations	36,134	17,418	18,716
(1) Simple LOC Index with Equal Weights[1]			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.023^{***}	0.015	0.022^{**}
	(0.008)	(0.010)	(0.011)
$(LOC_{P66}, LOC_{max}]$	0.036^{***}	0.030^{***}	0.038^{***}
	(0.008)	(0.011)	(0.011)
Observations	36,134	17,418	18,716
(2) Averaged LOC Factor[2]			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.026^{***}	0.019	0.033^{***}
	(0.008)	(0.012)	(0.012)
$(LOC_{P66}, LOC_{max}]$	0.046^{***}	0.042^{***}	0.041^{***}
	(0.009)	(0.013)	(0.013)
Observations	36,134	17,418	18,716
(3) LOC Factor 2005[3]			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.018^{**}	0.018	0.015
	(0.009)	(0.012)	(0.012)
$(LOC_{P66}, LOC_{max}]$	0.023^{**}	0.015	0.028^{**}
	(0.009)	(0.013)	(0.014)
Observations	34,088	16,418	17,670
All Controls	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
[1] LOC factor for each year is calculated using a simple average of all 8 items used and thus assuming equal weights of all items. The Index is then imputed using the same rule as in the Baseline.
[2] LOC factor is calculated as usual for each year but is imputed into all years using the average over the LOC observations from 2005 and 2010.
[3] LOC factor is calculated as usual for 2005 and is imputed forward into all years.

Table 4.A.5: Robustness Checks (Marginal Effects) - Outcome Variable

	All (1)	**Men** (2)	**Women** (3)
Baseline - Moderate of Heavy Drinking			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.026***	0.028***	0.018*
	(0.007)	(0.010)	(0.010)
$(LOC_{P66}, LOC_{max}]$	0.036***	0.031***	0.037***
	(0.008)	(0.011)	(0.012)
Observations	36,134	17,418	18,716
(1) Alternative Indicators			
(a) Abstainer *(vs. All Other Outcomes)*			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	-0.016***	-0.017***	-0.014*
	(0.005)	(0.006)	(0.007)
$(LOC_{P66}, LOC_{max}]$	-0.022***	-0.023***	-0.018**
	(0.005)	(0.007)	(0.008)
Observations	36,134	17,418	18,716
(b) Rare Drinkers *(vs. Abstainers)*			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.005	-0.010	0.015
	(0.010)	(0.014)	(0.013)
$(LOC_{P66}, LOC_{max}]$	0.005	0.006	0.003
	(0.011)	(0.016)	(0.015)
Observations	20,958	9,782	11,176
(2) Objective Measure 2016			
(a) High Consumption Amount or High Consumption Frequency			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.039***	0.039**	0.032**
	(0.012)	(0.019)	(0.015)
$(LOC_{P66}, LOC_{max}]$	0.036***	0.049**	0.030***
	(0.013)	(0.020)	(0.016)
Observations	9,244	4,280	4,964
(b) High Consumption Amount			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.003	0.017	-0.004
	(0.006)	(0.011)	(0.007)
$(LOC_{P66}, LOC_{max}]$	0.003	0.021*	-0.006
	(0.007)	(0.012)	(0.007)
Observations	9,244	4,280	4,964
(c) High Consumption Frequency			
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.030**	0.018	0.034**
	(0.012)	(0.018)	(0.015)
$(LOC_{P66}, LOC_{max}]$	0.033**	0.034*	0.034**
	(0.013)	(0.020)	(0.016)
Observations	9,244	4,280	4,964
All Controls	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

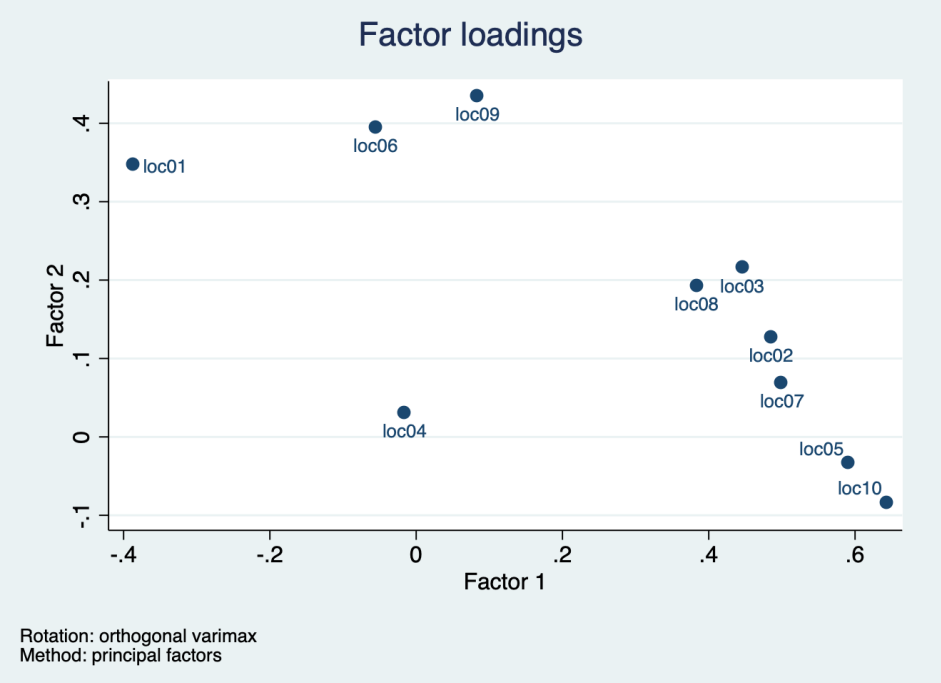
Table 4.A.6: Additional Results (Marginal Effects) - Compensating Health Behavior (Outcome: Moderate or Heavy Drinker)

	All (1)	**Men** (2)	**Women** (3)
Locus of Control Terciles (Ref.: $[LOC_{min}, LOC_{P33}]$)			
$(LOC_{P33}, LOC_{P66}]$	0.026***	0.027***	0.016
	(0.007)	(0.010)	(0.010)
$(LOC_{P66}, LOC_{max}]$	0.039***	0.031***	0.036***
	(0.008)	(0.011)	(0.012)
Smoking (Ref:: Non)			
Light	0.046***	0.039***	0.049***
	(0.009)	(0.012)	(0.012)
Heavy	0.058***	0.016	0.054***
	(0.010)	(0.013)	(0.017)
Healthy Diet (Ref:: Non)			
Moderate	-0.059***	-0.024*	-0.058**
	(0.012)	(0.013)	(0.023)
Strong	-0.142***	-0.125***	-0.116***
	(0.016)	(0.022)	(0.027)
Exercise (Ref:: Non)			
Moderate	0.041***	0.031***	0.042***
	(0.007)	(0.010)	(0.011)
Strong	0.063***	0.040***	0.087***
	(0.008)	(0.011)	(0.011)
Observations	36,035	17,377	18,658
All Controls	✓	✓	✓

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Clustered Standard Errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.B Supplementary Material

Figure 4.B.1: Factor Loadings of the LOC Variable



Source: SOEP, waves 2005, 2010 and 2015, version 33, doi:10.5684/soep.v33, own illustration.

Table 4.B.1: Descriptive Analysis - Willingness to Take Risk

	All			**Women**			**Men**		
	All (1)	Ext. (2)	Int. (3)	All (4)	Ext. (5)	Int. (6)	All (7)	Ext. (8)	Int. (9)
(1) General Willingness to take risk									
Continuous	4.51	4.28	4.74***	4.10	3.92	4.30***	4.95	4.70	5.17***
> Median	0.44	0.39	0.50***	0.34	0.30	0.38***	0.55	0.49	0.61***
(2) Willingness to Take Risk in Health									
Continuous	2.98	2.95	3.01***	2.64	2.69	2.60***	3.34	3.27	3.41***
> Median	0.43	0.43	0.44	0.36	0.38	0.35***	0.51	0.49	0.52***
(3) Deviation from General Willingness to Take Risk[1]									
Health	1.53	1.32	1.73***	1.46	1.23	1.70***	1.61	1.43	1.77***
Observations	36,134			18,716			17,418		

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Individuals are grouped into internals (Int.) and externals (Ext.) based on whether their LOC is lower/equal or higher than the median. Significance stars refer to significance level of t-test for mean equivalence between externals and internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[1] Deviation calculated as the difference between general and health-related willingness to take risk. Positive values indicating a lower health-related risk taking.

Table 4.B.2: Descriptive Analysis - Other Health Outcomes

	All			**Women**			**Men**		
	All (1)	Ext. (2)	Int. (3)	All (4)	Ext. (5)	Int. (6)	All (7)	Ext. (8)	Int. (9)
Smoking									
Non	0.70	0.69	0.72***	0.74	0.73	0.75***	0.67	0.64	0.69***
Light	0.17	0.16	0.17	0.17	0.17	0.17	0.16	0.16	0.16
Heavy	0.13	0.15	0.11***	0.09	0.10	0.08***	0.17	0.20	0.15***
Observations	36,035			18,658			17,377		
Healthy Diet									
Non	0.06	0.07	0.05***	0.03	0.04	0.03***	0.09	0.10	0.08***
Moderate	0.86	0.86	0.86	0.86	0.87	0.85***	0.86	0.85	0.86**
Strong	0.08	0.07	0.09***	0.11	0.09	0.12***	0.05	0.05	0.06***
Observations	36,134			18,716			17,418		
Exercise									
Non	0.34	0.39	0.30***	0.34	0.39	0.29***	0.34	0.40	0.30***
Moderate	0.28	0.28	0.27*	0.26	0.26	0.26	0.30	0.31	0.29**
Regular	0.38	0.32	0.43***	0.39	0.34	0.45***	0.36	0.30	0.41***
Observations	36,134			18,716			17,418		

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Individuals are grouped into internals (Int.) and externals (Ext.) based on whether their LOC is lower/equal or higher than the median. Significance stars refer to significance level of t-test for mean equivalence between externals and internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.B.3: Descriptive Analysis - Alcohol Consumption 2016

	All			**Women**			**Men**		
	All (1)	Ext. (2)	Int. (3)	All (4)	Ext. (5)	Int. (6)	All (7)	Ext. (8)	Int. (9)
How often do you drink alcohol?									
Every Day	0.07	0.07	0.07	0.03	0.03	0.04	0.11	0.11	0.11
4-6 days a week	0.08	0.07	0.09***	0.05	0.05	0.06*	0.12	0.10	0.13***
2-3 days a week	0.21	0.19	0.23***	0.17	0.15	0.19***	0.25	0.23	0.27***
2-4 days a month	0.24	0.24	0.25	0.24	0.24	0.25	0.24	0.23	0.25***
Once a month or less	0.24	0.25	0.23**	0.30	0.30	0.30	0.17	0.18	0.16**
Never	0.16	0.19	0.13***	0.20	0.23	0.18***	0.12	0.15	0.09***
Observations	9,246			4,966			4,280		
When you drink, how many drinks do you consume per day?									
1-2 drinks	0.69	0.70	0.68	0.79	0.78	0.79	0.59	0.60	0.58
3-4 drinks	0.23	0.22	0.23	0.17	0.17	0.17	0.29	0.29	0.30
5-6 drinks	0.05	0.05	0.05	0.03	0.03	0.03	0.08	0.08	0.08
7-9 drinks	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.03
10+ drinks	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01
Observations	7,751			3,964			3,787		

Source: SOEP, waves 2006, 2008, 2010, version 33, doi:10.5684/soep.v33, own calculations.
Notes: Individuals are grouped into internals (Int.) and externals (Ext.) based on whether their LOC is lower/equal or higher than the median. Significance stars refer to significance level of t-test for mean equivalence between externals and internals: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

CHAPTER 5

Biased by Success and Failure: How Unemployment Shapes Locus of Control

This chapter (pp. 144-170) is based on joint work with **Malte Preuß** and has been published in *Labour Economics*, Volume 53, pp. 63-74, https://doi.org/10.1016/j.labeco.2018.05.007.

Bibliography

Ahn, T. (2015). Locus of Control and Job Turnover. *Economic Inquiry*, **53** (2), 1350–1365.

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, **32** (4), 665–83.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, **115** (3), 715–753.

Alecke, B., Mitze, T. and Untiedt, G. (2010). Internal migration, regional labour market dynamics and implications for German East-West disparities: results from a Panel VAR. *Jahrbuch für Regionalwissenschaft*, **30** (2), 159–189.

Almlund, M., Duckworth, A. L., Heckman, J. and Kautz, T. (2011). *Personality Psychology and Economics*. Tech. rep., Institute for the Study of Labor (IZA), Bonn.

Andrienko, Y. and Guriev, S. (2004). Determinants of Interregional Mobility in Russia: Evidence from Panel Data. *Economics of Transition*, **12** (1), 1–27.

Andrisani, P. J. (1977). Internal-external attitudes, personal initiative, and the labor market experience of black and white men. *Journal of Human Resources*, **12** (3), 308–328.

Anger, S., Camehl, G. and Peter, F. H. (2017). Involuntary job loss and changes in personality traits. *Journal of Economic Psychology*, **60**, 71–91.

Angrist, J. D. (2002). How Do Sex Ratios Affect Marriage and Labor Markets ? Evidence from America' s Second Generation. *The Quarterly Journal of Economics*, **117** (3), 997–1038.

— and Evans, W. N. (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review*, **88** (3), 450–477.

— and Pischke, J. S. (2008). *Mostly harmless econometrics: an empiricist's companion*. Princeton University Press.

Anjomani, A. (2002). Regional growth and interstate migration. *Socio-Economic Planning Sciences*, **36** (4), 239–265.

Argys, L. M. and Rees, D. I. (2008). Searching for peer group effects: A test of the contagion hypothesis. *Review of Economics and Statistics*, **90** (3), 442–458.

Arni, P., Caliendo, M., Künn, S. and Zimmermann, K. F. (2014). The IZA evaluation dataset survey: a scientific use file. *IZA Journal of European Labor Studies*, **3** (1).

Arntz, M. (2010). What attracts human capital? Understanding the skill composition of interregional job matches in Germany. *Regional Studies*, **44** (4), 423–441.

—, Gregory, T. and Lehmer, F. (2011). *Unequal pay or unequal employment? What drives the self-selection of internal migrants in Germany*. Discussion Paper 11-074, ZEW, Mannheim.

Arulampalam, W., Gregg, P. and Gregory, M. (2001). Unemployment scarring. *The Economic Journal*, **111** (475), 577–584.

Ashenfelter, O. and Heckman, J. J. (1974). The Estimation of Income and Substitution Effects in a Model of Family Labor Supply. *Econometrica*, **42** (1), 73–85.

Autor, D. H., Katz, L. F. and Kearney, M. S. (2008). Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics*, **90** (2), 300–323.

Bailey, M. J. (2006). More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply. *Quarterly Journal of Economics*, **121** (1), 289–320.

Bauernschuster, S., Falck, O., Heblich, S., Suedekum, J. and Lameli, A. (2014). Why are educated and risk-loving persons more mobile across regions? *Journal of Economic Behavior & Organization*, **98** (6860), 56–69.

BBSR (2015). INKAR online. Federal Institute for Research on Building, Urban Affairs and Spatial Development, Online Resource, `http://www.inkar.de` (Accessed July 26, 2018).

Becker, A., Deckers, T., Dohmen, T., Falk, A. and Kosse, F. (2012). The Relationship Between Economic Preferences and Psychological Personality Measures. *Annual Review of Economics*, **4** (1), 453–478.

Becker, G. S. and Murphy, K. M. (1988). A Theory of Rational Addiction. *Journal of Political Economy*, **96** (4), 675–700.

Bennett, P., Norman, P., Murphy, S., Moore, L. and Tudor-Smith, C. (1998). Beliefs about alcohol, health locus of control, value for health and reported consumption in a representative population sample. *Health Education Research*, **13** (1), 25–32.

Berger, E. M. and Haywood, L. (2016). Locus of control and mothers' return to employment. *Journal of Human Capital*, **10** (4), 442–481.

Bertrand, M. (2010). New perspectives on gender. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4b, Elsevier, pp. 1543–1590.

—, Pan, J. and Kamenica, E. (2015). Gender Identity and Relative Income Within Households. *The Quarterly Journal of Economics*, **130** (2), 571–614.

Blanchard, O. J., Katz, L. F., Hall, R. E. and Eichengreen, B. (1992). Regional Evolutions. *Brookings Papers on Economic Activity,*, (1), 1–75.

Blau, F. D. (1998). Trends in the well-being of American women, 1970-1995. *Journal of Economic Literature*, **36**, 112–165.

— and Kahn, L. M. (2007). Changes in the labor supply behaviour of married women: 1980-2000. *Journal of Labor Economics*, **25** (3), 393–438.

— and — (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, **55** (3), 789–865.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, volume 3 edn., Amsterdam: Elsevier, pp. 1559–1695.

Boffetta, P. and Hashibe, M. (2006). Alcohol and cancer. *The Lancet Oncology*, **7** (2), 149–156.

Boneva, B. S. and Frieze, I. H. (2001). Toward a Concept of a Migrant Personality. *Journal of Social Issues*, **57** (3), 477–491.

Bonin, H., Constant, A., Tatsiramos, K. and Zimmermann, K. F. (2009). Native-migrant differences in risk attitudes. *Applied Economics Letters*, **16** (15), 1581–1586.

—, Eichhorst, W., Florman, C., Hansen, M. O., Skiöld, L., Stuhler, J., Tatsiramos, K., Thomasen, H. and Zimmermann, K. (2008). *Report No. 19: Geographic Mobility in the European Union: Optimising its Economic and Social Benefits*. IZA Research Reports 19, Institute for the Study of Labor (IZA).

Bono, J. E. and Judge, T. A. (2003). Core Self-Evaluations: A Review of the Trait and its Role in Job Satisfaction and Job Performance. *European Journal of Personality*, **17**, 5–18.

Booth-Kewley, S. and Vickers, R. R. (1994). Associations between Major Domains of Personality and Health Behavior. *Journal of Personality*, **62** (3), 281–298.

Borghans, L., Duckworth, A. L., Heckman, J. J. and ter Weel, B. (2008). The Economics and Psychology of Personality Traits. *Journal of Human Resources*, **43** (4), 972–1059.

—, Golsteyn, B. H., Heckman, J. J. and Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and Individual Differences*, **51** (3), 315–320.

Borjas, G. J. (1999). The economic analysis of immigration. In O. C. Ashenfelter and D. Card (eds.), *Handbook of labor economics*, volume 3, edn., North-Holland, pp. 1277–2097.

— (2000). *Labor Economics*. McGraw-Hill, 4th edn.

—, Bronars, S. G. and Trejo, S. J. (1992). Self-selection and internal migration in the United States. *Journal of Urban Economics*, **185**, 159–185.

Bosma, H., Schrijvers, C. and Mackenbach, J. P. (1999). Socioeconomic inequalities in mortality and importance of perceived control: cohort study. *British Medical Journal*, **319** (7223), 1469–1470.

Bouchard, T. J. and Loehlin, J. C. (2001). Genes, Evolution, and Personality. *Behavior Genetics*, **31** (3), 243–273.

Bound, J., Brown, C. and Mathiowetz, N. (2001). Measurement error in survey data. In J. J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, 5th edn., North-Holland, pp. 3705–3843.

Bowles, S., Gintis, H. and Osborne, M. (2001). Incentive-Enhancing Preferences: Personality, Behavior, and Earnings. *The American Economic Review*, **91** (2), 155–158.

Boyce, C., Wood, A. M., Daly, M. and Sedikides, C. (2015). Personality change following unemployment. *Journal of Applied Psychology*, **100** (4), 991–1011.

Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics*, **46** (4), 1171–1178.

Braun, M., Scott, J. and Alwin, D. F. (1994). Economic necessity or self-actualization? Attitudes toward women's labour-force participation in East and West Germany. *European Sociological Review*, **10** (1), 29–47.

Brewin, C. R. and Shapiro, D. A. (1984). Beyond locus of control: Attribution of responsibility for positive and negative outcomes. *British Journal of Psychology*, **75** (1), 43–49.

Bronars, S. G. and Grogger, J. (1994). The Economic Consequences of Unwed Motherhood : Using Twin Births as a Natural Experiment. *The American Economic Review*, **84** (5), 1141–1156.

Browning, M. and Heinesen, E. (2012). Effect of job loss due to plant closure on mortality and hospitalization. *Journal of Health Economics*, **31** (4), 599–616.

Buchinsky, M., Gotlibovski, C. and Lifshitz, O. (2014). Residential Location, Work Location, and Labor Market Outcomes of Immigrants in Israel. *Econometrica*, **82** (3), 995–1054.

Buddelmeyer, H. and Powdthavee, N. (2016). Can having internal locus of control insure against negative shocks? Psychological evidence from panel data. *Journal of Economic Behavior & Organization*, **122**, 88–109.

Buonanno, P. and Vanin, P. (2013). Bowling alone, drinking together. *Empirical Economics*, **44** (3), 1635–1672.

Caliendo, M., Cobb-Clark, D. A., Hennecke, J. and Uhlendorff, A. (2015a). Job Search, Locus of Control and Internal Migration. *IZA Discussion Paper*, **No. 9600**.

—, —, Seitz, H. and Uhlendorff, A. (2016). Locus of control and investment in training. *IZA Discussion Paper*, **No. 10406**.

—, — and Uhlendorff, A. (2015b). Locus of control and job search strategies. *Review of Economics and Statistics*, **97** (1), 88–103.

—, Fossen, F. and Kritikos, A. S. (2014). Personality characteristics and the decisions to become and stay self-employed. *Small Business Economics*, **42** (4), 787–814.

—, Künn, S. and Mahlstedt, R. (2017). The return to labor market mobility: An evaluation of relocation assistance for the unemployed. *Journal of Public Economics*, **148**, 136–151.

Carman, R. S. (1974). Internal-external locus of control, alcohol use and adjustment among high school students in rural communities. *Journal of Community Psycholog*, **2** (2), 129–133.

Carodso, A. R. and Morin, L.-P. (2018). Can Economic Pressure Overcome Social Norms? The Case of Female Labor Force Participation — IZA - Institute of Labor Economics. *IZA Discussion Paper Series*, **No. 11822**.

Cawley, J. and Ruhm, C. J. (2011). The Economics of Risky Health Behaviours. In M. V. Pauly, T. G. Mcguire and P. P. Barros (eds.), *Handbook of Health Economics*, *Chapter 3*, volume 2 edn., Elsevier B.V., pp. 95–199.

Chaloupka, F. (1991). Rational Addictive Behavior and Cigarette Smoking. *Journal of Political Economy*, **99** (4), 722–742.

Charles, K. K., Guryan, J. and Pan, J. (2018). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination. *NBER Working Paper Series*, **No. 24904**.

Chiteji, N. (2010). Time preference, noncognitive skills and well being across the life course : do noncognitive skills encourage healthy behavior? *The American Economic Review*, **100**, 200–204.

Christakis, N. A. and Fowler, J. H. (2007). The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine*, **357** (4), 370–379.

Clark, A. E. (2003). Unemployment as a Social Norm: Psychological Evidence from Panel Data. *Journal of Labor Economics*, **21** (2), 323–351.

—, Diener, E., Georgellis, Y. and Lucas, R. E. (2008). Lags and leads in life satisfaction: a test of the baseline hypothesis. *The Economic Journal*, **118** (529), 222–243.

Cobb-Clark, D. A. (2015). Locus of control and the labor market. *IZA Journal of Labor Economics*, **4** (1), 1–19.

—, Kassenboehmer, S. C. and Schurer, S. (2014). Healthy habits: The connection between diet, exercise, and locus of control. *Journal of Economic Behavior and Organization*, **98**, 1–28.

—, — and Sinning, M. G. (2016). Locus of control and savings. *Journal of Banking and Finance*, **73**, 113–130.

— and Schurer, S. (2011). The stability of big-five personality traits. *Economics Letters*, **115** (5943), 11 – 15.

— and — (2013). Two economists' musings on the stability of locus of control. *The Economic Journal*, **123** (570), 358–400.

— and Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, **18** (1), 1–13.

Cogan, J. (1980). Labor Supply with Costs of Labor Market Entry. In J. P. Smith (ed.), *Female Labor Supply: Theory and Estimation*, *Chapter 7*, Princeton University Press, pp. 327 – 364.

Coleman, M. and DeLeire, T. (2003). An economic model of locus of control and the human capital investment decision. *Journal of Human Resources*, **38** (3), 701–721.

Cook, P. J. and Moore, M. J. (2000). Alcohol. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1, *Chapter 30*, Elsevier B.V., pp. 1629–1673.

Corrao, G., Bagnardi, V., Zambon, A. and La Vecchia, C. (2004). A meta-analysis of alcohol consumption and the risk of 15 diseases. *Preventive Medicine*, **38** (5), 613–619.

Cortes, P. and Pan, J. (2018). Occupation and Gender. In S. L. Averett, L. M. Argys and S. D. H. Hoffman (eds.), *Oxford Handbook on the Economics of Women*, New York: Oxford University Press.

Costa, D. L. (2000). From Mill Town to Board Room: The Rise of Women's Paid Labor. *Journal of Economic Perspectives*, **14** (4), 101–122.

Cotter, D., Hermsen, J. M. and Vanneman, R. (2011). The End of the Gender Revolution? Gender Role Attitudes from 1977 to 2008. *American Journal of Sociology*, **117** (1), 259–289.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297–334.

Cseres-Gergely, Z. (2004). County to county migration and labour market conditions in Hungary between 1994 and 2002. *Journal for Labour Market Research, Institute for Employment Research*, **37** (4), 425–436.

Cull, A., Anderson, E. D., Campbell, S., Mackay, J., Smyth, E. and Steel, M. (1999). The impact of genetic counselling about breast cancer risk on women's risk perceptions and levels of distress. *British Journal of Cancer*, **79** (3-4), 501–508.

Cunha, F. and Heckman, J. J. (2007). The technology of skill formation. *The American Economic Review*, **97** (2), 31–47.

— and — (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, **43** (4), 738–782.

—, — and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, **78** (3), 883–931.

Cutler, D. M., Glaeser, E. L. and Shapiro, J. M. (2003). Why Have Americans Become More Obese? *Journal of Economic Perspectives*, **17** (3), 93–118.

— and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, **29** (1), 1–28.

DaVanzo, J. (1978). Does unemployment affect migration? Evidence from micro data. *The Review of Economics and Statistics*, pp. 504–514.

Davis, J. A. (1983). Does Authority Generalize? Locus of Control Perceptions in Anglo-American and Mexican-American Adolescents. *Political Psychology*, **4** (1), 101–120.

De Giorgi, G. and Pellizzari, M. (2009). Welfare migration in Europe. *Labour Economics*, **16** (4), 353–363.

den Berg, G. J. and Gorter, C. (1997). Job search and commuting time. *Journal of Business & Economic Statistics*, **15** (2), 269–281.

Devereux, P. J. (2004). Changes in Relative Wages and Family Labor Supply. *The Journal of Human Resources*, **39** (3), 696–722.

Dielman, T., Campanelli, P. C., Shope, J. T. and Butchart, A. T. (1987). Susceptibility to Peer Pressure, Self-Esteem, and Health Locus of Control as Correlates of Adolescent Substance Abuse. *Health Education Quarterly*, **14** (2), 207–221.

Diewald, M. (2007). Agency or constraint? The mutual shaping of control beliefs and working lives in East Germany after 1989. *International Journal of Psychology*, **42** (2), 77–84.

Duncan, G. J., Boisjoly, J., Kremer, M., Levy, D. M. and Eccles, J. (2005). Peer effects in drug use and sex among college students. *Journal of Abnormal Child Psychology*, **33** (3), 375–385.

— and Dunifon, R. (1998). "Soft-Skills" and Long-Run Labor Market Success. *Research in Labor Economics*, **17**, 123–149.

Dütsch, M. and Struck, O. (2014). Employment trajectories in Germany: Do firm characteristics and regional disparities matter? *Journal for Labour Market Research*, **47** (1-2), 107–127.

Ederveen, S., Bardsley, N. and Ederveen, S. (2003). *The influence of wage and unemployment differentials on labour mobility in the EU: A meta-analysis*. Research report, workpackage 3, ACCESSLAB.

Eissa, N. and Liebman, J. B. (1996). Labor Supply Response to the Earned Income Tax Credit. *The Quarterly Journal of Economics*, **111** (2), 605–637.

Eliasson, K., Lindgren, U. and Westerlund, O. (2003). Geographical Labour Mobility: Migration or Commuting? *Regional Studies*, **37** (8), 827–837.

Enchautegui, M. E. (1997). Welfare payments and other economic determinants of female migration. *Journal of Labor Economics*, **15** (2), 529–554.

Etzo, I. (2008). Internal migration: a review of the literature. *MPRA Paper*, (8783).

Evers, M., Mooij, R. and Vuuren, D. (2008). The wage elasticity of labour supply: A synthesis of empirical estimates. *Economist*, **156** (1), 25–43.

Falk, A. and Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, **362** (6412).

Farrell, P. and Fuchs, V. R. (1982). Schooling and Health: The Cigarette Connection. *Journal of Health Economics*, **1**, 217–230.

Fortin, N. M. (2015). Gender Role Attitudes and Women's Labor Market Participation: Opting-Out, AIDS, and the Persistent Appeal of Housewifery. *Annals of Economics and Statistics*, **117/118**, 379–401.

French, M. T. and Zarkin, G. A. (1995). Is moderate alcohol use related to wages? Evidence from four worksites. *Journal of Health Economics*, **14** (3), 319–344.

Frieze, I. H. and Li, M. Y. (2010). Mobility and personality. *The Psychology of Global Mobility*.

Frijling, B. D., Lobo, C. M., Keus, I. M., Jenks, K. M., Akkermans, R. P., Hulscher, M. E., Prins, A., Van Der Wouden, J. C. and Grol, R. P. (2004). Perceptions of cardiovascular risk among patients with hypertension or diabetes. *Patient Education and Counseling*, **52** (1), 47–53.

Fritsch, M. and Noseleit, F. (2013). Start-ups, long-and short-term survivors, and their contribution to employment growth. *Journal of Evolutionary Economics*, **23** (4), 719–733.

Fuchs-Schündeln, N. and Schündeln, M. (2009). Who stays, who goes, who returns? *Economics of Transition*, **17** (4), 703–738.

Furnham, A. (1986). Economic Locus of Control. *Human Relations*, **39** (1), 29–43.

— and Boo, H. C. (2011). A literature review of the anchoring effect. *Journal of Socio-Economics*, **40** (1), 35–42.

— and Greaves, N. (1994). Gender and locus of control correlates of body image dissatisfaction. *European Journal of Personality*, **8** (3), 183–200.

Gabriel, P. E. and Schmitz, S. (1995). Favorable self-selection and the internal migration of young white males in the United States. *Journal of Human Resources*, **30** (3), 460–471.

Gallo, W. T., Endrass, J., Bradley, E. H., Hell, D. and Kasl, S. V. (2003). The influence of internal control on the employment status of German workers. *Schmollers Jahr*, **123** (1), 71–82.

Giannetti, M. (2003). On the mechanics of migration decisions: skill complementarities and endogenous price differentials. *Journal of Development Economics*, **71** (2), 329–349.

Gintis, H. (2006). The Foundations of Behavior: The Beliefs, Preferences, and Constraints Model. *Biological Theory*, **1** (2), 123–127.

Giulietti, C. and Wahba, J. (2013). Welfare migration. In A. F. Constant and K. F. Zimmermann (eds.), *Handbook on the Economics of Migration*, vol. 2012-38, *Chapter 26*, Cheltenham, UK, and Northampton, USA: Edward Elgar, pp. p.489–504.

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C. and Schupp, J. (2018). The German Socio-Economic Panel (SOEP). *Journal of Economics and Statistics*.

Goldin, C. (1990). *Understanding the gender gap : an economic history of American women*. Oxford University Press, Inc.

— (2006). The Quiet Revolution that Transformed Women's Employment, Education, and Family. *The American Economic Review*, **96**, 1–21.

— and Katz, L. F. (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy*, **110** (4), 730–770.

Goldsmith, A. H., Veum, J. R. and Darity, W. (1997). The impact of psychological and human capital on wages. *Economic Inquiry*, **35** (4), 815–829.

Gottschalk, P. (2005). Can work alter welfare recipients' beliefs? *Journal of Policy Analysis and Management*, **24** (3), 485–498.

Graves, P. E. and Linneman, P. D. (1979). Household migration: Theoretical and empirical results. *Journal of urban economics*, **6** (3), 383–404.

Greenwood, J., Seshadri, A. and Yorukoglu, M. (2005). Engines of liberation. *Review of Economic Studies*, **72** (1), 109–133.

Greenwood, M. J. (1997). Internal migration in developed countries. In M. R. Rosenzweig and O. Stark (eds.), *Handbook of population and family economics*, Elservier Science, pp. 647–720.

— and Hunt, G. L. (2003). The Early History Of Migration Research. *International Regional Science Review*, **26** (1), 3–37.

Gronau, R. (1973). The Effect of Children on the Housewife's Value of Time. *Journal of Political Economy*, **81** (2), 168–199.

Grønbæk, M. (2009). The positive and negative health effects of alcohol- and the public health implications. *Journal of Internal Medicine*, **265** (4), 407–420.

Grønbæk, M., Becker, U., Johansen, D., Gottschau, A., Schnohr, P., Hein, H. O., Gorm, J. and Sørensen, MD, D. (2000). Type of Alcohol Consumed and Mortality from All Causes, Coronary Heart Disease, and Cancer. *Annals of Internal Medicine*, **133** (6), 411–419.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, **80** (2), 223–255.

— (2000). The Human Capital Model. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1, *Chapter 7*, Elsevier B.V., pp. 347–408.

Groves, M. O. (2005). How important is your personality? Labor market returns to personality for women in the US and UK. *Journal of Economic Psychology*, **26** (6), 827–841.

Gruber, J. and Köszegi, B. (2001). Is addiction "rational"? Theory and evidence. *Quarterly Journal of Economics*, **116** (4), 1261–1303.

Guglielminetti, E., Lalive, R., Ruh, P. and Wasmer, E. (2015). Spatial search strategies of job seekers and the role of unemployment insurance. *Working Paper, Sciences Po.*

Hainmueller, J. (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis*, **20** (1), 25–46.

— and Xu, Y. (2013). Ebalance: a Stata package for entropy balancing. *Journal of Statistical Software*, **54** (7), 1–18.

Hamaker, E. L., Nesselroade, J. R. and Molenaar, P. C. (2007). The integrated trait-state model. *Journal of Research in Personality*, **41** (2), 295–315.

Hamilton, V. and Hamilton, B. H. (1997). Alcohol and Earnings: Does Drinking Yield a Wage Premium? *The Canadian Journal of Economics*, **30** (1), 135–151.

Hansemark, O. C. (2003). Need for achievement, locus of control and the prediction of business start-ups: a longitudinal study. *Journal of Economic Psychology*, **24** (3), 301–319.

Harris, J. R. and Todaro, M. P. (1970). Migration , Unemployment and Developmnent : A Two-Sector Analysis. *The American Economic Review*, **60** (1), 126–142.

Hausman, J. A. (1980). The effect of wages, taxes, and fixed costs on women's labor force participation. *Journal of Public Economics*, **14** (2), 161–194.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, **19** (4), 451–464.

—, Moon, S. H., Pinto, R., Savelyev, P. A. and Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, **94** (1-2), 114–128.

—, Pinto, R. and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, **103** (6), 2052–2086.

— and Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*, **91** (2), 145–154.

—, Stixrud, J. and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, **24** (3), 411 – 482.

Heim, B. T. (2007). The Incredible Shrinking Elasticities: Married Female Labor Supply, 1978 2002. *J. Human Resources*, **XLII** (January 2007), 881–918.

Heineck, G. and Anger, S. (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, **17** (3), 535–546.

Hetschko, C., Knabe, A. and Schöb, R. (2014). Changing identity: retiring from unemployment. *The Economic Journal*, **124** (575), 149–166.

Heywood, J. S., Jirjahn, U. and Struewing, C. (2017). Locus of control and performance appraisal. *Journal of Economic Behavior & Organization*, **142**, 205–225.

Hicks, J. (1932). *The theory of wages*. London: Macmillan and Co Limited.

HILDA (2015). Household, Income and Labour Dynamics in Australia: data for years 2000-2014, file release 14.

Hill, R. (2011). *Teach internal locus of control: A positive psychology app*. Will To Power Press.

Hines, C., Koeppel, J. C. and Jacobs, K. W. (1974). Locus of Control as a variable in Geographic Mobility. *Perceptul and Motor Skills*, **39**, 1282.

Holt, C. L., Roth, D. L., Huang, J. and Clark, E. M. (2014). Gender Differences in the Roles of Religion and Locus of Control on Alcohol Use and Smoking Among African Americans. *Journal of Studies on Alcohol and Drugs*, **76** (3), 482–492.

Hoorens, V. and Buunk, B. P. (1993). Social comparison of health risks: locus of control, the person-positivity bias in favor of anonymous peers. *Journal of Applied Social Pscyhology*, **23** (4), 291–302.

Hunt, J. (2006). Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association*, **4** (5), 1014–1037.

Huttunen, K. and Kellokumpu, J. (2016). The effect of job displacement on couples' fertility decisions. *Journal of Labor Economics*, **34** (2), 403–442.

Infurna, F. J., Gerstorf, D., Ram, N., Schupp, J., Wagner, G. G. and Heckhausen, J. (2016). Maintaining perceived control with unemployment facilitates future adjustment. *Journal of Vocational Behavior*, **93**, 103–119.

International Labour Organization (2018). Labour Force. `http://www.ilo.org/global/statistics-and-databases/statistics-overview-and-topics/WCMS{_}470304/lang--en/index.htm`, accessed: 2018-05-07.

Jacobs, K. W. and Koeppel, J. C. (1974). Psychological correlates of the mobility decision. *Bulletin of the Psychonomic Society*, **3** (5A), 330–332.

Jaeger, B. D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., Bonin, H., Jaeger, D. A., Dohmen, T. and Falk, A. (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics*, **92** (August), 684–689.

Jahoda, M. (1981). Work, employment, and unemployment: Values, theories, and approaches in social research. *American Psychologist*, **36**, 184–191.

James, W. H. (1957). *Internal versus external control of reinforcement as a basis variable in learning theory*. Ph.D. thesis.

Janz, N. K. and Becker, M. H. (1984). The Health Belief Model: A Decade Later. *Health Education & Behavior*, **11** (1), 1–47.

John, O. P. and Srivastava, S. (1999). The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. Second edi edn., New York: The Guildford Press, pp. 102 – 138.

Jokela, M. (2009). Personality predicts migration within and between US states. *Journal of Research in Personality*, **43**, 79–83.

Jones, A. S. and Richmond, D. W. (2006). Causal effects of alcoholism on earnings: Estimates from the NLSY. *Health Economics*, **15** (8), 849–871.

Judge, T. A., Erez, A., Bono, J. E. and Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, **83** (3), 693–710.

—, —, — and — (2003). The core self-evaluations scale: Development of a measure. *Personnel Psychology*, **56** (2), 303–331.

— and Hurst, C. (2007). Capitalizing on One's Advantages: Role of Core Self-Evaluations. *Journal of Applied Psychology*, **92** (5), 1212–1227.

Juhn, C. and Murphy, K. M. (1997). Wage Inequality and Family Labor Supply. *Journal of Labor Economics*, **15** (1), 72–97.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, **47** (March), 263–291.

Källmén, H. (2000). Manifest anxiety, general self-efficacy and locus of control as determinants of personal and general risk perception. *Journal of Risk Research*, **3** (2), 111–120.

Kassenboehmer, S. C. and Haisken-DeNew, J. P. (2009). You're fired! The causal negative effect of entry unemployment on life satisfaction. *The Economic Journal*, **119** (536), 448–462.

Katz, L., Kling, J. and Liebman, J. (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, **116** (2), 607–654.

Keil, U., Liese, A., Filipiak, B., Swales, J. D. and Grobbee, D. E. (1998). Alcohol, Blood Pressure and Hypertension. In D. J. Chadwick and J. A. Goode (eds.), *Novartis Found Symposium*, vol. 216, Novartis Foundation, pp. 125–158.

Kenkel, D. S. (1991). Health Behavior, Health Knowledge, and Schooling. *Journal of Political Economy*, **99** (2), 287–305.

Kennan, J. and Walker, J. R. (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica*, **79** (1), 211–251.

Kesavayuth, D., Ko, K. M. and Zikos, V. (2018). Locus of control and financial risk attitudes. *Economic Modelling*, **72**, 122–131.

Killingsworth, M. R. and Heckman, J. J. (1986). Female labor supply: A survey. In O. C. Ashenfelter and R. L. Layard (eds.), *Handbook of Labor Economics*, vol. 1, *Chapter 2*, Amsterdam: Elsevier Science Publishers BV, pp. 103–204.

Kling, J. R., Liebman, J. B. and Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, **75** (1), 83–119.

Knabe, A., Schöb, R. and Weimann, J. (2016). Partnership , Gender Roles and the Well-Being Cost of Unemployment. *Social Indicators Research*, **129** (3), 1255–1275.

Koenig, C. S. and Cunningham, W. R. (2001). Adulthood relocation: implications for personality, future orientation, and social partner choices. *Experimental aging research*, **27** (2), 197–213.

Konle-Seidl, R., Eichhorst, W. and Grienberger-Zingerle, M. (2007). Activation Policies in Germany: From Status Protection to Basic Income Support. *IAB Discussion Paper, Beiträge zum wissenschaftlichen Dialog aus dem Institut für Arbeitsmarkt- und Berufsforschung*, **06/2007**.

Kreiner, C. T. and Tranæs, T. (2005). Optimal workfare with voluntary and involuntary unemployment. *Scandinavian Journal of Economics*, **107** (3), 459–474.

Kunze, L. and Suppa, N. (2017). Bowling alone or bowling at all? The effect of unemployment on social participation. *Journal of Economic Behavior & Organization*, **133**, 213–235.

Lefcourt, H. M. (1966). Internal versus external control of reinforcement: A review. *Psychological Bulletin*, **65** (4), 206–220.

— (1982). *Locus of Control: Current Trends in Theory and Research*. Oxford, England: Lawrence Erlbaum Associaties, Publishers, 2nd edn.

— (1992). Durability and impact of the locus of control construct. *Psychological Bulletin*, **112** (3), 411–414.

Legerski, E., Cornwall, M. and O'Neil, B. (2006). Changing locus of control: steelworkers adjusting to forced unemployment. *Social Forces*, **84** (3), 1521–1537.

Leibenstein, H. (2006). Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *The Quarterly Journal of Economics*, **64** (2), 183–207.

Leifman, H., Kühlhorn, E., Allebeck, P., Andréasson, S. and Romelsjö, A. (1995). Abstinence in late adolescence-Antecedents to and covariates of a sober lifestyle and its consequences. *Social Science and Medicine*, **41** (1), 113–121.

Lekfuangfu, W. N., Powdthavee, N., Warrinnier, N. and Cornaglia, F. (2018). Locus of Control and its Intergenerational Implications for Early Childhood Skill Formation. *The Economic Journal*, **128** (608), 298–329.

Lemos-Giráldez, S. and Fidalgo-Aliste, A. M. (1997). Personality dispositions and health-related habits and attitudes: A cross-sectional study. *European Journal of Personality*, **11** (3), 197–209.

Levy, M. B. and Wadycki, W. J. (1974). Education and the Decision to Migrate: An Econometric Analysis of Migration in Venezuela. *Econometrica*, **42** (2), 377–388.

Long, J. S. and Freese, J. (2014). *Regression Models for Categorical Dependent Variables Using Stata*. College Station, TX: Stata Press, 3rd edn.

Lucas, R. E. B. (1997). Internal migration in developing countries. *Handbook of population and family economics*, **1**, 721–798.

Ludwig, J., Duncan, G. J. and Pinkston, J. C. (2005). Housing mobility programs and economic self-sufficiency. *Journal of Public Economics*, **89** (1), 131–156.

— and Kling, J. (2007). Is Crime Contagious? *The Journal of Law and Economics*, **50** (3), 491–518.

Lundberg, S. (1988). Labor supply of husbands and wives: A simultaneous equations approach. *The Review of Economics and Statistics*, **70** (2), 224–235.

— (2012). Personality and Marital Surplus Personality and Marital Surplus. *IZA Journal of Labor Economics*, **1** (3), 1–21.

Lundborg, P. (2006). Having the wrong friends? Peer effects in adolescent substance use. *Journal of Health Economics*, **25** (2), 214–233.

Lutgen, V. and Van der Linden, B. (2015). Regional equilibrium unemployment theory at the age of the Internet. *Regional Science and Urban Economics*, **53**, 50–67.

Macdonald, Z. and Shields, M. A. (2001). The impact of alcohol consumption on occupational attainment in England. *Economica*, **68** (271), 427–453.

— and — (2004). Does problem drinking affect employment? Evidence from England. *Health Economics*, **13** (2), 139–155.

Manski, C. F. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives*, **14** (3), 115–136.

Marchiori, L. and Schumacher, I. (2011). When nature rebels: international migration, climate change, and inequality. *Journal of Population Economics*, **24** (2), 569–600.

Marcus, J. (2014). Does job loss make you smoke and gain weight? *Journal of Health Economics*, **32** (324), 546–558.

McAdams, D. P. and Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, **61** (3), 204–217.

McCrae, R. R. and Costa Jr, P. T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Personality Psychology*, **3** (6), 173–175.

— and Costa Jr, P. T. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins and L. A. Pervin (eds.), *Handbook of Personality: Theory and Research*, New York: Guilford Press, pp. 159–181.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In D. McFadden and P. Zarembka (eds.), *Frontiers in econometrics*, *Chapter Fo*, New York: Academic Press, pp. 105–142.

McGee, A. D. (2015). How the perception of control influences unemployed job search. *ILR Review*, **68** (1), 184–211.

— and McGee, P. (2016). Search, effort, and locus of control. *Journal of Economic Behavior and Organization*, **126** (2), 89–101.

McGrattan, E. R. and Rogerson, R. (2008). Changes in the Distribution of Family Hours Worked Since 1950. In P. Rupert (ed.), *Frontiers of Family Economics*, *Chapter 3*, Amsterdam: Emerald Group Publishing Limited, pp. 115 – 138.

Mendolia, S. and Walker, I. (2014a). The effect of noncognitive traits on health behaviours in adolescence. *Health economics*, **23** (9), 1146–1158.

— and — (2014b). The effect of personality traits on subject choice and performance in high school: Evidence from an English cohort. *Economics of Education Review*, **43**, 47–65.

Merikangas, K. R., Mehta, R. L., Molnar, B. E., Walters, E. E., Swendsen, J. D., Aguilar-Gaziola, S., Bijl, R., Borges, G., Caraveo-Anduaga, J. J., Dewit, D. J., Kolody, B., Vega, W. A., Wittchen, H. U. and Kessler, R. C. (1998). Comorbidity of substance use disorders with mood and anxiety disorders: Results of the international Consortium in Psychiatric Epidemiology. *Addictive Behaviors*, **23** (6), 893–907.

Mincer, J. (1962). Labor Force Participation of Married Women: A Study of Labor Supply. In National Bureau Committee for Economic Research (ed.), *Aspects of Labor Economics*, Princeton University Press, pp. 63–106.

— (1985). Inter-country comparisons of labor force trends and of related developments: An overview. *Journal of Labor Economics*, **3** (1), 1–32.

Moretti, E. (2011). Local Labor Markets. In O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, volume 4b edn., North-Holland, pp. 1237–1313.

Morrison, P. A. and Wheeler, J. P. (1978). The image of "elsewhere" in the American tradition of migration. In William H. McNeill and Ruth S. Adams (ed.), *Human migration: patterns and policies*, Indiana University Press Bloomington Ind. United States 1978., pp. 75–84.

Mueller, G. and Plug, E. (2006). Estimating the effect of personality on male and female earnings. *Industrial and Labor Relations Review*, **60** (1), 3–22.

Mullahy, J. and Sindelar, J. (1996). Employment, unemployment, and problem drinking. *Journal of Health Economics*, **15** (4), 409–434.

— and Sindelar, J. L. (1993). Alcoholism, Work, and Income. *Journal of Labor Economics*, **11** (3), 494–520.

Nenov, P. T. (2015). Regional reallocation and housing markets in a model of frictional migration. *Review of Economic Dynamics*, **18**, 863–880.

Ng, T. W. H., Sorensen, K. L. and Eby, L. T. (2006). Locus of control at work: a meta-analysis. *Journal of Organizational Behavior*, **27** (8), 1057–1087.

Niebuhr, A., Granato, N., Haas, A. and Hamann, S. (2012). Does labour mobility reduce disparities between regional labour markets in Germany? *Regional Studies*, **46** (7), 841–858.

NLSY79 (2019). NLSY79 Appendix 21: Attitudinal Scales. `www.nlsinfo.org/`, accessed: 2019-03-20.

Nolte, H., Weischer, C., Wilkesmann, U., Maetzel, J. and Tegethoff, H. (1997). Kontrolleinstellungen zum Leben und zur Zukunft: Auswertung eines neuen sozialpsychologischen Itemblocks im sozio-ökonomischen Panel. *Diskussionspapiere aus der Fakiltät für Sozialwissenschaft, Ruhr-Universität Bochum*, **97-06**.

Nyhus, E. K. and Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, **26** (3), 363–384.

Offerhaus, J. (2013). The type to train? Impacts of personality characteristics on further training participation. *SOEPpapers*, **531**.

Oppenheimer, V. K. (1970). *The female labor force in the United States; demographic and economic factors governing its growth and changing composition*. Westport, CT: Greenwood Press.

Ours, J. V. (2004). A pint a day raises a man's pay; but smoking blows that gain away. *Journal of health economics*, **23** (5), 863–886.

Owen, D. and Green, A. (1992). Migration Processes and Patterns. Volume 1. Research Progress and Prospects. In A. Champion and A. Fielding (eds.), *Migration processes and patterns*, vol. 1, *Migration*, London: Belhaven: Belhaven Press: Kings Lynn, pp. 17–38.

Pareto, V. (1906). *Manual of Political Economy: A Critical and Variorum Edition*. Oxford University Press, english edn.

Parikh, A. and van Leuvensteijn, M. (2002). *Internal Migration in Regions of Germany: A Panel Data Analysis*. Working Paper 12, European Network of Economic Policy Research Institutes.

Pearlin, L. I. and Schooler, C. (1978). The Structure of Coping. *Journal of Health and Social Behavior*, **19** (1), 2–21.

Peter, F. H. and Spiess, C. K. (2016). Family instability and locus of control in adolescence. *The B.E. Journal of Economic Analysis & Policy*, **16** (3), 1439–1471.

Peters, B. L. and Stringham, E. (2006). No booze? You may lose: Why drinkers earn more money than nondrinkers. *Journal of Labor Research*, **27** (3), 411–421.

Peters, R., Peters, J., Warner, J., Beckett, N. and Bulpitt, C. (2008). Alcohol, dementia and cognitive decline in the elderly: A systematic review. *Age and Ageing*, **37** (5), 505–512.

Phares, E. J. (1957). Expectancy changes in skill and chance situations. *Journal of Abnormal and Social Psychology*, **54** (3), 339–342.

Piatek, R. and Pinger, P. (2010). Maintaining (Locus of) Control ? Assessing the Impact of Locus of Control on Education Decisions and Wages. *SOEPpapers*, **338**.

— and — (2016). Maintaining (locus of) control? Data combination for the identification and inference of factor structure models. *Journal of Applied Econometrics*, **31** (4), 734–755.

Pinger, P., Schäfer, S. and Schumacher, H. (2018). Locus of control and consistent investment choices. *Journal of Behavioral and Experimental Economics*, **75** (11537), 66–75.

Plunkett, H. R. and Buehner, M. J. (2007). The relation of general and specific locus of control to intertemporal monetary choice. *Personality and Individual Differences*, **42** (7), 1233–1242.

Preuss, M. and Hennecke, J. (2018). Biased by success and failure: How unemployment shapes locus of control. *Labour Economics*, **53**, 63–74.

Puhani, P. a. (2001). Labour Mobility: An Adjustment Mechanism in Euroland? Empirical Evidence for Western Germany, France and Italy. *German Economic Review*, **2** (2), 127–140.

Rachlin, H. (1997). Four teleological theories of addiction. *Psychonomic Bulletin and Review*, **4** (4), 462–473.

Rainer, H. and Siedler, T. (2009). The role of social networks in determining migration and labour market outcomes. *Economics of Transition*, **17** (4), 739–767.

Rauch, A. and Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology*, **16** (4), 353–385.

Rehm, J., Greenfield, T. and Rogers, J. (2001). Average volume of alcohol consumption, patterns of drinking, and all-cause mortality: Results from the US national alcohol survey. *American Journal of Epidemiology*, **153** (1), 64–71.

Reimers, C. (1985). Cultural Differences in Labor Force Participation Among Married Women. *the American Economic Review*, **75** (2), 251–255.

Renaud, S., Guéguen, R., Schenker, J. and D'Houtaud, A. (1998). Alcohol and Mortality in Middle-Aged Men from Eastern France. *Epidemiology.*, **9** (2), 184–188.

Richter, D., Metzing, M., Weinhardt, M. and Schupp, J. (2013). *SOEP Scales Manual*. Tech. rep., SOEP Survey Papers 138: Series C.

Roback, J. (1982). Wages , Rents , and the Quality of Life. *Journal of Political Economy*, **90** (6), 1257–1278.

Roberts, B. W. (2009). Back to the future: personality and assessment and personality development. *Journal of Research in Personality*, **43** (2), 137–145.

— and DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychological bulletin*, **126** (1), 3–25.

—, Harms, P., Smith, J. L., Wood, D. and Webb, M. (2006a). Using Multiple Methods in Personality Psychology. In M. Eid and E. Diener (eds.), *Handbook of Multimethod Measurement in Psychology*, *Chapter 22*, American Psychological Association, pp. 321–335.

—, Jackson, J. J., Duckworth, A. L. and Von Culin, K. (2011). Personality Measurement and Assessment in Large Panel Surveys. *Forum for Health Economics & Policy*, **14** (3).

— and Mroczek, D. (2008). Personality Trait Change in Adulthood. *Current Directions in Psychological Science*, **17** (1), 31–35.

—, Walton, K. E. and Viechtbauer, W. (2006b). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin*, **132** (1), 1–25.

—, Wood, D. and Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins and L. A. Pervin (eds.), *Handbook of Personality: Theory and Research*, vol. 3, New York: Guilford Press, pp. 375–398.

Ronksley, P. E., Brien, S. E., Turner, B. J., Mukamal, K. J. and Ghali, W. A. (2011). Association of alcohol consumption with selected cardiovascular disease outcomes: A systematic review and meta-analysis. *British Medical Journal*, **342** (7795), 479.

Roosma, F., van Oorschot, W. and Gelissen, J. (2016). The Achilles' heel of welfare state legitimacy: perceptions of overuse and underuse of social benefits in Europe. *Journal of European Public Policy*, **23** (2), 177–196.

Rosenbaum, M. (1980). A schedule for assessing self-control behaviors: Preliminary findings. *Behavior Therapy*, **11** (1), 109–121.

Rosenfeld, R. A., Trappe, H. and Gornick, J. C. (2004). Gender and Work in Germany: Before and After Reunification. *Annual Review of Sociology*, **30** (1), 103–124.

Rosengren, A., Hawken, S., Ounpuu, S., Sliwa, P., Zubaid, M., Almahmeed, W., Ngu Blackett, K., Sitthi-Amorn, C., Sato, H. and Yusuf, P. (2004). Association of psychosocial risk factors with risk of acute myocardial infarction in 11 119 cases and 13 648 controls from 52 countries (the Interheart study): Case-control study. *Lancet*, **364** (9438), 953–962.

Ross, C. E., Mirowsky, J. and Huber, J. (1983). Dividing work, sharing work, and in-between: marriage patterns and depression. *American sociological review*, **48** (6), 809–823.

Rotter, J. B. (1954). *Social Learning and Clinical Psychology*. Englewood Cliffs, NJ, US: Prentice-Hall, Inc.

— (1966). Generalized Expectancies for Internal Versus External Control of Reinforcement. *Psychological Monographs: General and Applied*, **80** (1), 1–28.

Rouwendal, J. (1999). Spatial job search and commuting distances. *Regional Science and Urban Economics*, **29** (4), 491–517.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, **3** (2), 135–146.

Ryon, H. S. and Gleason, M. E. (2014). The Role of Locus of Control in Daily Life. *Personality and Social Psychology Bulletin*, **40** (1), 121–131.

Saks, R. E. and Wozniak, A. (2011). Labor Reallocation over the Business Cycle: New Evidence from Internal Migration. *Journal of Labor Economics*, **29** (4), 697–739.

Salamanca, N., De Grip, A., Fouarge, D. and Montizaan, R. (2016). Locus of Control and Investment in Risky Assets. *IZA Discussion Paper Series*, **No. 10407**.

Sampson, R. and Wooldredge, J. (1986). Evidence that high crime rates encourage migration away from central cities. *Sociology and Social Research*.

Schmitz, H. (2011). Why are the unemployed in worse health? The causal effect of unemployment on health. *Labour Economics*, **18** (1), 71–78.

Schnitzlein, D. D. and Stephani, J. (2016). Locus of control and low-wage mobility. *Journal of Economic Psychology*, **53**, 164–177.

Schöb, R. (2013). Unemployment and identity. *CESifo Economic Studies*, **59** (1), 149–180.

Schurer, S. (2017). Bouncing back from health shocks: Locus of control and labor supply. *Journal of Economic Behavior and Organization*, **133**, 1–20.

Semykina, A. and Linz, S. J. (2007). Gender differences in personality and earnings: Evidence from Russia. *Journal of Economic Psychology*, **28** (3), 387–410.

Simon, H. A. (1982). *Models of bounded rationality: Economic analysis and public policy*, vol. 1. The MIT Press.

Sjaastad, L. A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy*, **70** (5), 80–93.

Sjoberg, L. (2000). Factors in Risk Perception. *Risk Analysis*, **20** (1), 1–11.

Skandalis, D. (2019). *Breaking News: The Role of Information in Job Search and Matching*. Working paper.

Skinner, E. A. (1996). A Guide to Constructs of Control. *Journal of Personality and Social Psychology*, **71** (3), 549–570.

Slovic, P. (1992). Perception of Risk: Reflections on the Psychometric Paradigm. In D. Golding and S. Krimsky (eds.), *Social Theories of Risk*, pp. 117–152.

Smith, J. P. and Ward, M. P. (1985). Time-Series Growth in the Female Labor Force. *Journal of Labor Economics*, **3** (2), 59–90.

SOEP (2017). Socio Economic Panel Data: Data for years 1984-2016, version 33, Doi: 10.5684/soep.v33.

Soto, C. J. and Tackett, J. L. (2015). Personality Traits in Childhood and Adolescence: Structure, Development, and Outcomes. *Current Directions in Psychological Science*, **24** (5), 358–362.

Specht, J., Egloff, B. and Schmukle, S. C. (2013). Everything under control? The effects of age, gender, and education on trajectories of perceived control in a nationally representative German sample. *Developmental Psychology*, **49** (2), 353–364.

Steptoe, A. and Wardler, J. (2001). Locus of control and health behaviour revisited: A multivariate analysis of young adults from 18 countries. *British Journal of Psychology*, **92**, 659–672.

Stillwell, J., Rees, P. and Duke-Williams, O. (1996). Population Migration in the European Union. In A. C. P. Rees J. Stillwell and M. Kupiszewski (eds.), *Population migration in the European Union*, *Migration*, John Wiley, pp. 275–307.

Stürmer, T., Hasselbach, P. and Amelang, M. (2006). Personality, lifestyle, and risk of cardiovascular disease and cancer: Follow-up of population based cohort. *British Medical Journal*, **332** (7554), 1359–1362.

Suranovic, S. M., Goldfarb, R. S. and Leonard, T. C. (1999). An economic theory of cigarette addiction. *Journal of Health Economics*, **18** (1), 1–29.

Thakker, K. D. (1998). An overview of health risks and benefits of alcohol consumption. *Alcoholism: Clinical and Experimental Research*, **22** (7), 285–298.

Thaler, R. H. (2015). *Misbehaving: The Making of Behavioral Economics*. New York: W.W. Norton & Company.

— and Shefrin, H. M. (1981). An Economic Theory of Self-Control. *Journal of Political Economy*, **89** (2), 392–406.

Thomas, D. S. (1938). *Research Memorandum on Migration Differentials*. New York: Social Science Research Council.

Todaro, M. P. (1969). A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *The American Economic Review*, **59** (1), 138–148.

Toney, M. B., Pitcher, B. L. and Stinner, W. F. (1985). Geographic Mobility and Locus of Control. *The Journal of psychology*, **119** (4), 361–368.

Topel, R. H. (1986). Local Labor Markets. *Journal of Political Economy*, **94** (S3), 111–143.

Touraine, A. and Ragazzi, O. (1961). Ouvriers d'origine agricole. *Aux Editions du Seuil*.

Treyz, G. I., Rickman, D. S., Hunt, G. L. and Greenwood, M. J. (1993). The Dynamics of U.S. Internal Migration. *The Review of Economics and Statistics*, **75** (2), 209–214.

Uhlendorff, A. (2004). Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, **56** (2), 279–303.

Valletta, R. G. (2013). House lock and structural unemployment. *Labour Economics*, **25**, 86–97.

Van Den Berg, G. J., Bergemann, A. H. and Caliendo, M. (2009). The Effect of Active Labor Market Programs on Not-Yet Treated Unemployed Individuals. *Journal of the European Economic Association*, **7** (2-3), 606–616.

— and Van Der Klaauw, B. (2006). Counseling and Monitoring of Unemployed Workers: Theory and Evidence from a Controlled Social Experiment. *International Economic Review*, **47** (3), 895–936.

van Ommeren, J., van den Berg, G. J. and Gorter, C. (2000a). Estimating the marginal willingness to pay for commuting. *Journal of Regional Science*, **40** (3), 541–563.

van Ommeren, J. N., Rietveld, P. and Nijkamp, P. (1999). Job moving, residential moving, and commuting: a search perspective. *Journal of Urban Economics*, **46** (2), 230–253.

—, — and — (2000b). Job mobility, residential mobility and commuting: A theoretical analysis using search theory. *The Annals of Regional Science*, **34** (2), 213–232.

Viscusi, W. (1990). Do Smokers Underestimate Risks? *Journal of Political Economy*, **98** (6), 1253–1269.

Vollrath, M., Knoch, D. and Cassano, L. (1999). Personality, risky health behaviour, and perceived susceptibility to health risks. *European Journal of Personality*, **13** (1), 39–50.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.

Wagner, G. G., Frick, J. and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP): scope, evolution and enhancements. *Schmollers Jahrbuch*, **127** (1), 139–170.

Wallston, B. S. and Wallston, K. A. (1978). Locus of Control and Health: A Review of Literature. *Health Education Monographs*, **6**, 107–117.

—, — and DeVellis, R. (1978). Development of the Multidimensional Health Locus of Control (MHLC) Scales. *Health Education Monographs*, **6** (2), 160–170.

—, —, Kaplan, G. D. and Maides, S. A. (1976). Development and Validation of the Health Locus of Control (HLC) Scale. *Journal of Consulting and Clinical Psychology*, **44** (4), 580–585.

Wang, L.-Y., Kick, E., Fraser, J. and Burns, T. J. (1999). Status attainment in America: The roles of locus of control and self-esteem in educational and occupational outcomes. *Sociological Spectrum*, **19** (3), 281–298.

Weinberg, B. A. (2000). Computer use and the demand for female workers. *Industrial and Labor Relations Review*, **53** (2), 290–308.

Weinstein, N. D. (1984). Why It Won't Happen to Me: Perceptions of Risk. *Health Psychology*, **3** (5), 431–457.

— (1987). Unrealistic optimism about susceptibility to health problems. *Journal of Behavioral Medicine*, **10** (5), 481–500.

Wichern, F. and Nowicki, S. (1976). Independence training practices and locus of control orientation in children and adolescents. *Developmental Psychology*, **12** (1), 77.

Wichert, L. and Pohlmeier, W. (2010). Female labor force participation and the big five, ZEW Discussion Paper 10-003. *ZEW Discussion Paper*, **No. 10-003**.

Wijatkowski, S., Forgays, D. G., Wrzesniewski, K. and Gorski, T. (1990). Smoking behavior and personality characteristics in polish adolescents. *International Journal of the Addictions*, **25** (4), 363–373.

Winefield, A., Winefield, H. and Tiggemann, M. (1991). A longitudinal study of the psychological effects of unemployment and unsatisfactory employment on young adults. *Journal of Applied Psychology*, **76**, 424–431.

World Health Organization (2016). Global Health Observatory data repository.

Wozniak, A. (2010). Are College Graduates More Responsive to Labor Market Opportunities? *Journal of Human Resources*, **45** (4), 944–970.

Zax, J. S. (1991). The substitution between moves and quits. *The Economic Journal*, **101** (409), 1510–1521.

— (1994). When is a move a migration? *Regional Science and Urban Economics*, **24** (3), 341–360.

— and Kain, J. F. (1991). Commutes, quits, and moves. *Journal of Urban Economics*, **29** (2), 153–165.

Ziebarth, N. R. and Grabka, M. M. (2009). In vino pecunia? The association between beverage-specific drinking behavior and wages. *Journal of Labor Research*, **30** (3), 219–244.

Abstract

Investigating the psychological black box behind individual economic decision making is, without a doubt, one of the most prevalent concerns in recent empirical microeconomics. This is based on the urge of modern behavioral economics to provide the stochastic idiosyncratic shocks in standard economic models with meaningful content. While traditional approaches largely characterized individual deviations from the prescriptions of the standard models as individual error or irrational behavior, an important differentiation between normative and positive economics has emerged again over the past decades. Behavioral approaches in modern empirical microeconomics are largely driven by the goal of drawing a picture of human behavior which is as accurate as possible while still being informed by the prescription from utility maximization, full rationality and perfect information as the core ideas of economic research. Especially the growing availability of large microdata sources such as in longitudinal household panel studies has tremendously supported this scientific movement. This data regularly includes important self-reported information on inherent attributes such as preferences, expectations, concerns and personality traits which have a high potential of explaining large parts of the deviations which have previously been labeled as stochastic shocks and idiosyncratic errors.

This is the point at which also this doctoral thesis lines up. The present thesis contains four studies that investigate the relationship of inherent personality traits with individual behavior and economic outcomes. Concretely, the studies address the domains female labor force participation, labor market mobility, drinking behavior and unemployment. All studies make use of data from the Socio-Economic Panel (SOEP), a large-scale longitudinal study of representative households in Germany. Importantly, it is goal of all studies to not only empirically identify a relationship but also to pin down the knowledge gains and implications of the findings for traditional economic decision making models by carefully discussion channels and mechanism within the framework of rational choice theory.

The unifying element of all four studies is the focus on one specific personality trait within this context: the individual perception of control or locus of control (LOC). LOC is characterized as a *"generalized attitude, belief, or expectancy regarding the nature of the causal relationship between one's own behavior and its consequences"* (Rotter, 1966, p.2) and

describes whether individuals believe in the causal effects of their own efforts and abilities on their lives' outcomes. While individuals with an internal LOC belief that their own efforts and abilities are rewarded in their future life's outcome, individuals with an external LOC attribute life's outcomes mainly to luck, chance, fate or other people. The trait is already widely used within empirical microeconomics and has been found to have important effects on economic behavior and decision making especially on the labor market due to its effect on individual expectations and motivation. Within the economic and psychological literature, LOC is thus largely assumed to be an important determinant of individual efforts, actions and investments. All these existing ideas are supported and enhanced by the considerations and findings of the present thesis.

Chapter 2 initiates the discussion by analyzing the implications of LOC for female labor force participation. This is done by incorporating a vector of individual attributes into a woman's utility and budget function which captures LOC. The vector of attributes is argued to affect the preferences for leisure, home production and childcare as well as participation but also the subjective expectations about monetary returns to participation. Based on the psychological literature which makes an important connection between LOC and independence considerations, internal women are expected to gain higher direct non-monetary utility from participation as they not only derive utility from the consumption level as an outcome of participation but also from the fact that they themselves had control over the generation of it. The empirical analysis supports these theoretical considerations. Internal women are found to have a significantly higher probability of being available for market production, which also translates into higher employment probabilities at the extensive margin. These effects are additionally found to be highly heterogenous with respect to underlying monetary incentives for participation and home production as well as prevalent social norms for working.

In a quite similar manner, Chapter 3 discusses the role of LOC for regional labor market mobility within Germany. The empirical analysis identifies a distinct positive effect of an internal LOC on the general self-reported willingness to move as well as the probability of moving between regions. These findings support the theoretical considerations about LOC being an important determinant of geographic distant job search in line with earlier findings on the positive relationship between LOC and job seekers' subjective beliefs about returns to search effort. Internal job seekers are predicted to search across larger geographic areas and migrate more frequently if job search occurs before migration.

A prove that the importance of LOC for decision making cross the boarders of labor economics is provided within Chapter 4. The chapter is devoted to the question of whether LOC is also able to explain alcohol consumption as an important domain of risky health behavior. By making use of the unique information available on drinking behavior within the SOEP, the study identifies a strongly positive effect of an internal LOC on moderate as well as heavy drinking which is comparable to effect of traditional preference parameters such as

risk aversion and time preferences. Based on a number of ancillary estimations, the paper suggests and discusses two likely mechanisms for this relationship. On the one hand, internal individuals are expected to put a higher effort into the investment in social and occupational networks and via this have a higher likelihood to drink as the attendance of social gatherings is often inextricably linked with alcohol consumption. On the other hand, as uncertainty and the perception of risks play a crucial role for individual considerations about alcohol consumption, internals consume more alcohol on average as they believe in or overestimate their personal ability to cope with or prevent the negative consequences of consumption.

Eventually, Chapter 5 importantly contributes to the value all parts of this thesis adds to the body of literature on behavioral consequences of LOC by carefully discussion the stability of the trait and thus potential problems with reverse causality in the three other studies as well as in the existing literature in general. In order to access the stability of LOC, the study investigates the reaction of reported LOC to an exogenous unemployment shock. Reassuringly, the empirical analysis finds no long-lasting effects of job losses due to plant closures on LOC and thus cannot reject the common assumption of its stability during adulthood. Hence, concerns with respect to endogeneity of LOC within the models and thus reverse causality can be alleviated. Nevertheless, the study identifies an important temporary deviation in the measurement of LOC during periods of unemployment and therefore concludes that the reported LOC is affected by unemployment likely due to a situation-specific state effect. Rather than being concerned with reverse causality, the future literature thus needs to carefully consider potential measurement issues within LOC. The role of those findings for the studies in the previous chapters are carefully discussed in the introductory remarks as well as the chapters themselves.

Kurzzusammenfassung

Die Erforschung der psychologischen Blackbox, welche ökonomischer Entscheidungsfindung zugrunde liegt, ist ohne Zweifel eines der vorherrschenden Anliegen der jüngeren empirischen Mikroökonomik. Dies beruht vor allem auf dem Drang der modernen Verhaltensökonomik, die stochastischen idiosynkratischen Schocks in den ökonomischen Standardmodellen mit sinnstiftendem Inhalt zu füllen. Während traditionelle Ansätze individuelle Abweichungen von den Gesetzen der Standardmodelle größtenteils als Fehler und irrationales Verhalten einordneten, kam in den vergangenen Jahrzehnten eine wichtige Unterscheidung zwischen normativer und positiver Ökonomik erneut zum Vorschein. Verhaltensökonomische Ansätze in der modernen empirischen Mikroökonomik verfolgen zu einem großen Teil das Ziel, ein möglichst vollständiges Bild von menschlichem Verhalten zu zeichnen, und sich dennoch an den Gesetzen der Nutzenmaximierung, der vollkommenen Rationalität und der vollständigen Informationen als den Kernen ökonomischen Denkens zu orientieren. Vor allem die wachsende Verfügbarkeit von großen Mikrodatenquellen, wie unter anderem großangelegte Längsschnittstudien, hat diese wissenschaftliche Bewegung entscheidend vorangetrieben. Beruhend auf regelmäßigen Selbstbeurteilungen beinhalten diese Daten wichtige Informationen zu inhärenten Eigenschaften wie Präferenzen, Erwartungen, Anliegen und Persönlichkeitsmerkmale. Alle diese Eigenschaften haben ein hohes Potential große Teile dessen zu erklären, was bisher als stochastische Schocks und idiosynkratische Fehler bezeichnet wurde.

An diesem Punkt reiht sich auch die vorliegende Dissertation ein. Die Arbeit beinhaltet vier Studien, welche alle das Ziel verfolgen, den Zusammenhang zwischen inhärenten Persönlichkeitsmerkmalen und individuellem Verhalten sowie ökonomischen Ergebnissen zu untersuchen. Ganz konkret beschäftigen sich die Studien mit den Themen weibliche Arbeitsmarktbeteiligung, Arbeitsmarktmobilität, Alkoholkonsum sowie Arbeitslosigkeit. Hierbei nutzen alle Studien die Daten des Sozio-oekonomischen Panels (SOEP), einer großangelegten Langzeitbefragung repräsentativer Haushalte in Deutschland. Hierbei ist das Ziel der Studien nicht nur, einen empirischen Zusammenhang zu schätzen, sondern den Erkenntnisgewinn und die Implikationen der identifizierten Effekte für traditionelle ökonomische Entscheidungsmodelle herauszuarbeiten. Dafür werden jeweils potentielle Kanäle und Mechanismen innerhalb des Rahmens der Rational-Choice-Theorie sorgfältig diskutiert.

Das verbindende Element aller vier Studien ist der Fokus auf ein spezifisches Persönlichkeitsmerkmal in diesem Kontext: die individuelle Kontrollüberzeugung, der "locus of control" (LOC). LOC wird bezeichnet als eine generalisierte Einstellung oder Erwartung bezüglich der Art des kausalen Zusammenhangs zwischen eigenem Verhalten und seinen Konsequenzen (vgl. Rotter, 1966, S.2) und beschreibt, ob Individuen an den kausalen Effekt von eigenen Anstrengungen und Fähigkeiten auf Ergebnisse im Leben glauben. Während Individuen mit einer internalen Kontrollüberzeugung daran glauben, dass eigene Anstrengungen und Fähigkeiten mit zukünftigen Ereignissen belohnt werden, schreiben Individuen mit einer externalen Kontrollüberzeugung Lebensereignisse vor allem Glück, Schicksal, Zufall und anderen Personen zu. Diese Persönlichkeits-eigenschaft ist in der empirischen Mikroökonomik bereits weit verbreitet und es existieren eine Reihe von Studien, welche ihr eine wichtige Bedeutung für ökonomisches Verhalten vor allem auf dem Arbeitsmarkt zuschreiben. Dies ist vor allem begründet in der Rolle von LOC für Erwartungen sowie Motivation von Individuen. In der ökonomischen sowie psychologischen Literatur wird daher weitestgehend angenommen, dass Kontrollüberzeugung eine wichtige Determinante individueller Anstrengung, Aktivität und Investitionsentscheidung ist. Alle diese existierenden Ideen und Erkenntnisse werden von den Studien der vorliegenden Arbeit unterstützt und erweitert.

Die Diskussion wird in Kapitel 2 mit der Analyse der Implikationen von LOC für weibliche Arbeitsmarktbeteiligung eingeleitet. Hierzu wird ein Vektor mit individuellen Eigenschaften in die standardmäßige Nutzenfunktion und Budgetgleichung der Frauen eingeführt. Dieser Vektor beinhaltet auch die Kontrollüberzeugung und es wird angenommen, dass er die Präferenzen der Frauen für Freizeit, Hausarbeit und Kinderbetreuung sowie Arbeitsmarktbeteiligung, aber auch die subjektiven Erwartungen der Frauen über monetäre Vorteile von Partizipation entscheidend beeinflusst. Basierend auf der psychologischen Literatur, welche eine wichtige Verbindung zwischen Kontrollüberzeugung und individuellen Unabhängigkeitsbestrebungen macht, wird angenommen, dass Frauen mit einer internalen Kontrollüberzeugung höheren nicht-monetären Nutzen aus Partizipation ziehen, da sie nicht nur Nutzen aus dem Konsumlevel als Folge von Partizipation ziehen, sondern auch aus der Tatsache, dass sie die Kontrolle über das Erreichen dieses Konsumlevels hatten. Die empirische Analyse untermauert diese theoretischen Überlegungen. Die Ergebnisse zeigen, dass Frauen mit einer internalen Kontrollüberzeugung eine signifikant höhere Wahrscheinlichkeit haben für die Arbeitsmarktbeschäftigung verfügbar zu sein. Diese erhöhte Bereitschaft übersetzt sich außerdem auch in höhere extensive Beschäftigungswahrscheinlichkeiten. Zusätz-lich zeigen die empirischen Ergebnisse deutlich eine Heterogenität der Effekte bezüglich der zu Grunde liegenden monetären Arbeitsanreize sowie der vorherrschenden sozialen Normen zur Arbeitsmarktbeteiligung.

In einer sehr ähnlichen Weise diskutiert Kapitel 3 die Rolle von Kontrollüberzeugung für regionale Arbeitsmarktmobilität in Deutschland. Die empirische Analyse identifiziert einen

deutlichen positiven Effekt von internaler Kontrollüberzeugung auf die individuelle Bereitschaft zum Umzug sowie die tatsächliche Wahrscheinlichkeit eines Umzugs zwischen Arbeitsmarktregionen. Diese Ergebnisse bestätigen die angenommenen theoretischen Überlegungen. Hierbei wird, auf Basis existierender Forschung zum Effekt von LOC auf die subjektive Effektivität von Arbeitsplatzsuche, angenommen, dass LOC eine entscheidende Determinante geografisch weitangelegter Stellensuche ist. Internale Arbeitssuchende haben eine höhere Wahrscheinlichkeit über größere geografische Distanz nach Stellen zu suchen und haben deshalb eine höhere interregionale Umzugswahrscheinlichkeit unter der Voraussetzung, dass die Stellensuche vor der Migration stattfindet.

Ein Beweis dafür, dass die Bedeutung von Kontrollüberzeugung auch die Grenzen der Arbeitsmarktökonomik überschreitet, wird in Kapitel 4 gebracht. Das Kapitel beschäftigt sich mit der Frage, ob LOC ebenfalls in der Lage ist, Heterogenität im Alkoholkonsum als wichtigen Aspekt des individuellen riskanten Gesundheitsverhalten zu erklären. Unter Nutzung der Informationen zu Alkoholkonsum, welche im SOEP verfügbar sind, identifiziert die Studie starke positive Effekte von internaler Kontrollüberzeugung auf moderates und starkes Trinkverhalten. Der Effekt ist vergleichbar mit dem von traditionellen Determinanten von riskantem Gesundheitsverhalten wie u.a. der Risikobereitschaft oder der Zeitpräferenzen der Individuen. Basierend auf den empirischen Ergebnissen eine Reihe zusätzlicher Schätzungen, diskutiert das Papier zwei mögliche Mechanismen für diesen Zusammenhang. Einerseits wird erwartet, dass internale Individuen größeren Aufwand im Ausbau ihrer sozialen und beruflichen Netzwerke betreiben. Da die Teilnahme an gesellschaftlichen Zusammenkünften oft unumgänglich mit dem Konsum von Alkohol verbunden ist, haben diese Individuen eine erhöhte Wahrscheinlichkeit mit Alkohol in Berührung zu kommen. Andererseits ist Alkoholkonsum, mehr als anderes Gesundheitsverhalten, mit einem hohen Maß an Unsicherheit bezüglich der potentiellen negativen Folgen verbunden. Aus diesem Grund konsumieren internale Individuen möglicherweise mehr Alkohol, da sie an ihre persönliche Fähigkeiten glauben mit den negativen Folgen umzugehen oder diese zu verhindern.

Schlussendlich erhöht vor allem Kapitel 5 den Mehrwert, den alle diese Studien zur Literatur über die verhaltensmäßigen Konsequenzen von Kontrollüberzeugung beitragen. Dies liegt begründet in der vorherrschenden Diskussion zur Stabilität dieses Persönlichkeitsmerkmals und hiermit der potentiellen Probleme mit umgekehrter Kausalität und Endogenität, welche zentraler Untersuchungsgegenstand der Studie in Kapitel 5 ist. Um die Stabilität von Kontrollüberzeugung zu bewerten, analysiert die Studie den Effekt von exogenen Arbeitsplatzverlusten auf LOC. Als positives Signal für die Validität der empirischen Ergebnisse der anderen drei Studien sowie der Literatur im Allgemeinen, findet die Untersuchung keine langanhaltenden Effekte von Arbeitsplatzverlusten durch Firmenschließungen auf die Kontrollüberzeugung und damit auch keinen Hinweis auf ein Problem in der allem zu Grunde liegenden Annahme der Stabilität von Persönlichkeitsmerkmalen im Erwachsenenalter. Sorgen über

mögliche umgekehrte Kausalität oder Endogenität in den Modellen können daher weitestgehend gemildert werden. Allerdings identifiziert die Studie eine wichtige temporäre Abweichung in der Messung von LOC während Arbeitslosigkeitsperioden und kommt aus diesem Grund zu dem Schluss, dass die von Befragten angegebene Kontrollüberzeugung von einem wichtigen situationsspezifischen Statuseffekt beeinflusst wird. An Stelle von Bedenken zu umgekehrter Kausalität sollte die zukünftige Literatur aus diesem Grund vor allem mögliche Probleme in der Messung von Persönlichkeitsmerkmalen berücksichtigen und diskutieren. Die Implikationen dieser Ergebnisse für die Schätzungen in den anderen Kapiteln werden in den einleitenden Bemerkungen sowie in den Studien selbst vorsichtig diskutiert und analysiert.