UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

LABOR MARKET DYNAMICS IN THE SHORT-RUN IN CHILE.

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

CAROLINA ANDREA GALLEGUILLOS SAN MARTÍN

PROFESOR GUÍA:
BENJAMÍN VILLENA ROLDÁN

MIEMBROS DE LA COMISIÓN:
ALEJANDRA MIZALA SALCES
ANDREA REPETTO LISBOA

Este trabajo ha sido parcialmente financiado por

Beca CONICYT folio 22140419
Proyecto FONDECYT Regular 1151479
Proyecto CONICYT PIA SOC 1402

SANTIAGO DE CHILE
2018

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: CAROLINA ANDREA GALLEGUILLOS SAN MARTÍN
FECHA: 2018
PROF. GUÍA: SR. BENJAMÍN VILLENA ROLDÁN

LABOR MARKET DYNAMICS IN THE SHORT-RUN IN CHILE.

Esta tesis analiza las dinámicas de corto plazo del mercado laboral chileno entre 2003 y 2013. Estudiar dicho periodo requiere hacer una consideración importante: después que Chile se convirtiera en miembro de la OCDE en 2010, la Encuesta Nacional de Empleo (ENE) fue modificada de acuerdo a los parámetros estadísticos de dicha organización, pasando a ser conocida también como Nueva Encuesta Nacional de Empleo (NENE). Estos cambios implicaron pasar de estudiar los estados laborales desde una perspectiva de las labores principales durante "la mayor parte de la semana" (ENE), a estudiarlos considerando labores remuneradas al menos una hora en la semana de referencia —criterio de una hora en la definición del estado laboral— (NENE). Este cambio se tradujo en una pérdida de comparabilidad entre ambas mediciones, por lo que los estudios sobre el mercado laboral considerando tanto la ENE como la NENE no son posibles.

Este trabajo propone un método para superar esta dificultad por medio de la creación de contrafactuales de la probabilidad de estar en un estado laboral. Para obtener dichas probabilidades se estiman dos modelos: un primer modelo en función de características personales (individuales y del hogar), y un segundo modelo usando características personales y condiciones agregadas del mercado laboral segmentado por edad y región de residencia (efectos agregados). Además, la estimación se hace separadamente por sexo. Tomando los parámetros obtenidos en la medición antigua de la encuesta (ENE), se construyen los contrafactuales de la probabilidad de estar empleado, desempleado y fuera del mercado laboral (o inactivo) proyectando dichos parámetros en los datos post-2010 (periodo NENE), generando una serie de datos continua.

Un resultado básico de los contrafactuales es que, si el método de medición de los estados laborales no hubiera cambiado, las dinámicas del mercado laboral habrían permanecido relativamente estables entre 2003 y 2013, pese a la Crisis Subprime. La recuperación *post* crisis sería más gradual que la observada en los resultados de la NENE, siguiendo, además, tendencias similares previas a la Crisis. Un tema importante que surge es la pérdida de ciclicidad de los resultados de la NENE respecto de la ENE.

Las diferencias entre las probabilidades observadas y contrafactuales dan algunas pistas para analizar los resultados de la NENE. Primero, las fluctuaciones en las dinámicas del mercado laboral son principalmente atribuibles al cambio en la medición de los estados laborales. Segundo, los análisis de la evolución de los estados laborales deben considerar la participación en el mercado, o la relación entre empleo e inactividad. Desde una perspectiva más general, es claro que las fluctuaciones en las condiciones del mercado laboral dependen de los ciclos económicos y de las herramientas usadas para medir dichas condiciones, en particular, son evidentes los efectos de usar el criterio de una hora.

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: CAROLINA ANDREA GALLEGUILLOS SAN MARTÍN
FECHA: 2018
PROF. GUÍA: SR. BENJAMÍN VILLENA ROLDÁN

LABOR MARKET DYNAMICS IN THE SHORT-RUN IN CHILE.

This thesis analyzes the short-run dynamics of the Chilean labor market for the period 2003-2013. Studying this length of time must consider a critical data issue: after becoming an OECD member in 2010, the National Employment Survey (NES) was modified to meet the standards of this organization, called since then as the New National Employment Survey (NNES). It transforms from delimiting labor force status considering the economical activity the most important part of reference week to defining if surveyed worked at least one hour for pay or profit —one-hour criterion in the definition of employment—. This consideration affects the comparability of data across the surveys, making the study of labor market dynamics using both the NES and the NNES impractical.

This work proposes a method to achieve this goal using counterfactuals of the probability of being in a certain labor force status. There are two models to estimate these probabilities: the first model estimates probabilities as a function of personal characteristics and households conditions; and the second model considers personal characteristics, household conditions, and adds aggregate conditions of labor markets segmented by age and region of living (*aggregate effects*). Also, models are separately estimated by sex. Using parameters estimated in previous method of measurement (NES), they are generated counterfactual measures of employment, unemployment, and being out of labor market (or inactivity) using forecasting in the period after 2010 (NNES period), linking survey data.

A basic result of counterfactuals is that if the measurement method of the labor force status had not changed, labor market dynamics would have been steady between 2003 and 2013, despite Subprime Mortgage Crisis. Thus, recovery after crisis had been more gradual than observed NNES results indicate, following similar trends prior to the crisis. An important issue is the lack of cyclicality of the NNES results comparing to the NES.

Differences between observed data and counterfactuals give some clues to analyze the NNES results. First, fluctuations in labor market dynamics since 2010 are attributable, in an important share, to changes in the measurement of the labor force status. Second, any explanation of labor force status evolution should include the participation in the labor force or the relation between employment and inactivity as a much more precise thermometer of the labor market performance. From an overall perspective, it is clear that fluctuations in the labor market conditions depend on the business cycle and the tools used to measure them. Particularly, there are apparent the effects of using the one-hour criterion in the definition of employment.

"The problem of structure and agency has rightly come to be seen as the basic issue in modern social theory."
Margaret Scotford Archer

"Collective action is a means of power, a means by which individuals can more fully realize their individual values."
Kenneth Joseph Arrow

"The long run is a misleading guide to current affairs. In the long run we are all dead."
John Maynard Keynes

Agradecimientos

A mi familia, por el apoyo y el cariño entregado desde siempre. A mi mamá, Patricia, por su humor negro y su optimismo a toda prueba que me resulta imposible. A mi abuelo, Humberto, quien habría corregido la tesis con su estilo elegante y pulcro.

A mi profesor guía Benjamín Villena, por toda su apoyo y paciencia así como entusiasmo para que este trabajo saliera adelante. A las profesoras Alejandra Mizala y Andrea Repetto, por su disposición y por sus comentarios para mejorar este trabajo. A la profesora Marcela Valenzuela, por sus consejos prácticos para trabajar con forecasting.

A todas las personas increíbles que conocí en este (largo) proceso. Y a lo mucho que aprendí con (y de) ellas, de economía, de la vida y de mí misma (aunque las dos últimas probablemente sean parte del síndrome de Estocolmo).

A los profesores del magíster, por transmitir su entusiasmo y rigurosidad en el quehacer de la economía. En particular, al profesor Patricio Valenzuela, por la confianza en el trabajo y sus consejos.

A mis amigos, la familia que uno escoge. En particular a los amigos del MagCEA, compañeros de todas las batallas. A Pablo Carvajal, Nincen Figueroa, Eduardo Toro y José Miguel Alvarado, por su amistad y por su valiosa ayuda para esta tesis. A Diego Huerta, por su apoyo en la implementación de los modelos. A Martín Castillo por ser él. A las Valentinas y la resistencia femenina. A Gerardo Rojas por su optimismo, y a Camilo Levenier, por su apoyo cotidiano y referencias constantes a Los Simpson. A Camila Mella, por su primicidad y su minucioso apoyo en mi intento de escribir en inglés.

A la Sociología por enseñarme a hacer preguntas. A la Economía por darme herramientas para intentar responderlas.

A Felipe, por impulsarme a este desafío. Por ser y estar.

Contents

List of Tables

List of Figures

Introduction

Labor market is the place where workers and jobs (employers) interact with each other. The interaction between the supply and the demand side gives a dynamic character to the market. Due to its role, labor market has two important characteristics: first, it is a fundamental institution where individual and collective interests interact; and, second, it rarely is a perfectly competitive.

The labor market is a fundamental institution because of its central role in distributing rewards that are tied to one's social and economic status in society. It is important to consider that there are two effective mechanisms for contributing to the distribution of wealth. One of them corresponds to public policies, understood as the money paid to people who receive welfare benefits like pensions and tax credits. Another one is the labor market; which, in the context of a developing country, is the main mechanism through which growth and economic development are passed on to individuals (Sehnbruch, 2006).

The economy-wide labor supply is given by adding the work choices made by each person in the population, which has been explained by the neoclassical model of labor-leisure choice. In particular, it concerns the individual labor supply derived from several issues concerning "household production" (the preparation of meals, housekeeping, the raising of children) and household conditions alike (personal wealth, spousal income, family situation). So, any discussion of labor market must consider two things: on the one hand, the well-being of people and the macroeconomic equilibrium; and, on the other, how designing public policies without altering markets and their general equilibrium.

From above, it is important to remark that the labor market is imperfectly competitive. In fact, the labor market is not like the commodity market, where buyers and sellers meet and can trade any homogenous good because goods are non-differentiated and have easily verifiable quality (as in the cases of precious metals, oil, and electricity). Additionally, in the labor market there are rents associated with any given job, so the total surplus is positive (therefore, it does not meet the market-clearing condition). On the contrary, in the labor market, wages are a rent splitting device. Thereby, how labor market operates is closely related to the level of income inequality existing in any given society (Boeri & Van Ours, 2008).

A far-reaching consequence of the imperfect competition in the labor market (in opposition to the commodity market) is the importance of measuring labor market conditions and their assessment. Indeed, policy makers are desperately searching for the best methods to measure the health and the strength of the labor market because different data points, and even the same ones from month to month, may be confusing signals at the moment. One fact everyone should be able to agree on is that the official unemployment rate does not even attempt to measure the strength or health of the labor market, at least not as central criterion —despite the relationship between unemployment and government popularity leads to a policy focus on unemployment rates—.

Because the labor market depends on local and global institutions as well as economic growth and its cycles, its composition has been changing in the last fifty years —the period between 1950 and 1973 is also known as the "Golden Age", due to world growth rates was higher than at any other time in history—. One of the most dramatic changes has been the increasing labor force participation of women, particularly among married women (Blau & Kahn, 2013; Ehrenberg & Smith, 2009).

Another important change has been the increasing importance of human capital and the non-human capital, as well (Acemoglu, 2008). Human capital includes accumulated investments that allow workers embodying a set of skills that can be "rented out" to employers as education, job training, and migration. In the long-term, according to this theory, education increases productivity. In the short-term, (potential) employees send a signal about their ability level to employers by acquiring education credentials, and so, gaining access to certain occupations in the labor market, segmenting labor market.

Chile has not escaped aforementioned trends. Between 1988 and 1998, the Chilean economy experienced one of the highest growth periods in its history, with rates of around 6 percent. However, this growth period was interrupted in 1999 by the the Asian crisis, when Chile experienced its first recession since 1982-1983. The Chilean economy only recovered from this crisis in 2004, when returning to the pre-crisis growth rates. Nevertheless, it is important to remark that the key problem was the persistent high unemployment rate (Bravo, Ferrada, & Landerretche, 2005; Cowan, Micco, Mizala, Pagés, & Romaguera, 2005).

Nowadays, the Chilean economic outlook seems to be quite the opposite. According to data provided by the Central Bank of Chile, economic growth has decelerated sharply from 6.1 percent in 2011 to 1.6 percent in 2016. Nevertheless, and despite the slower growth, unemployment rate has remained stable due to an increase in self-employment (although the proportion of salaried employees fell).

Abovementioned dynamics illustrates that the rise and fall of economic growth (or business cycle) do not explain economic fluctuations. On the opposite, it is merely an aid to understand them. Therefore, labor market is the place to look for an understanding of the depth and persistence of recessions (Hall, 2005). Despite that the the unemployment rate is the most widely cited measure of labor market conditions, it is a crude estimate with several imperfections (Ehrenberg & Smith, 2009).

Data availability and its processing are important issues that stand out from this work. The National Employment Survey, principal source of labor data in Chile, was modified in 2010 to fulfill international requirements to join to OECD. Thus, changes introduced a new method, which introduced comparability problems with data produced before.

To solve the last point, this work proposes a method to correct this discontinuity based on estimating counterfactuals of labor force status if the change of methodology did not have happened. Counterfactuals consider the probability of being in the labor force status by gender. Probabilities are estimated for each observation using probit models, considering personal characteristics, household conditions, and structural labor market characteristics considering sociodemographic groups (defined by age, education, and region of residence). Afterwards, individual probabilities are added into market probabilities by gender.

A comparison of these two different methods makes possible to study labor market dynamics in recent years without discontinuities in data, and then, making to shed some light into the development of the Chilean labour market in recent decades. Therefore, it is possible to question: When did definitely begin labor market recovery from the Subprime mortgage crisis? To what extent measurement changes have affected labor trends data? Because labor is the most abundant factor of production, it is fair to say that any country's well-being in the long run depends heavily on the willingness of its people to work. But, Is still true in short-run?

In this regard, this thesis aims to answer this question: *What are the main dynamics of the Chilean labor market between 2003 and 2013?*. Therefore, the main objective is *to characterize the main labor market fluctuations in Chile during this period.* Additionally, this research also has other aims, such as: In first place, *to describe the labor market evolution in Chile between 2003 and 2013*; and, secondly, *to determine the impact of the change in the measurement method in the labor market fluctuations throughout this period.*

Several hypotheses support this work. First, the evolution of any labor market depends on business cycles in the short and mid-term, and the Chilean is not an exemption. This is what happened in the late nineties (Bravo et al., 2005; Cowan et al., 2005; Landerretche, 2007), and, since that there have been no sizable structural changes in the economy, it is the most reasonable to suppose this. Second, labor market dynamics are more cyclical than new measurement illustrates if we consider "comparable" methods. Additionally, the unemployment rate post-Subprime mortgage crisis remained fairly stable and its recovery has been weaker than official data illustrates. Thus, post-2010 fluctuations are attributable to changes in the methodology through which labor force status is measured (Marcel & Naudon, 2016). Third, aggregate effects are suitable instruments to add market conditions into an individual prediction of labor force status. These effects illustrate that the framework to understand how do people participate in the labor market, both to who are inside (as employed or unemployed) or who stay out (as inactive).

The outline is as follows. chapter 1 examines the contribution to literature to place this thesis within a wider context. This chapter reviews about measurements of labor market conditions as well as empirical challenges that countries should consider in measuring official statistics, dynamics of labor market in Chile from late nineties, and sociodemographic trends that affect labor market. chapter 2 discusses the characteristics of the data sources used in this work, its advantages and disadvantages. chapter 3 describes the proposed method to deal with data comparability problems, and then its validation. chapter 4 presents the main findings of Chilean labor market dynamics using the proposed method. Finally, the thesis concludes with a discussion of the role of measurement of labor force conditions and their evolution in the short-run, and also proposes directions for further research.

Chapter 1

Contribution to Literature

The labor market has undergone deep transformations in the last decades. As a result, international organizations have been urged to redesign and redefine indicators used to measure and characterize this market in order to adapt to new realities and actors. However, *What is the best course of action to handle these changes and the need for long-series data to study labor market?* From a more general perspective, *How can we have got updated data to fulfill current needs?*

This thesis introduces a proposal to deal with aforementioned dilemmas. Particularly, this chapter is organized as follows. The first section includes a brief overview of the international methods to measure labor force status and analyzes some countries where these methods have been amended. The second section describes the labor market performance in Chile in recent years, specifically, during the period when the measurement change was implemented. The third section reviews some demographic trends that affect labor market functioning, specifically, labor force status.

1.1 The Measurement of Labor Market Conditions

Aggregate economic variables in capitalist economies experience repeated fluctuations about their long-term growth paths. There are recurrent sequences of expansions and contractions in the aggregate economic activity. These downward and upward movements in longterm trends —measured by the Gross Domestic Product (GDP)— are known as business cycles (Barro & Sala-i Martin, 2004).

Business cycle fluctuations also involve labor market dynamics. Henceforth, a *labor market* is a market where a quantity of labor services, L, corresponding to tasks specified in an unfilled assignment or job description (vacant job), is offered in exchange for a price or remuneration, called wage, w (Boeri & Van Ours, 2008). Not all labor services offered by an individual are paid, but in order to get a job —at least for National Accounts—, there must be an exchange of a labor service for a wage.

Analyzing the labor market must consider how the demand and the supply of labor are related. On the demand side are employers, whose decisions about labor hiring are influenced by conditions in all markets (capital, labor, and product markets). On the supply side are workers and potential workers, whose decisions about where, and whether to work (or not) must take into account their other options about how to spend their time.

It is important to remark that any discussion of labor market must consider the conditions of people who are supplying their labor and the macroeconomic equilibrium. Equally important are the concepts used to measure labor market conditions, which denote personal and market conditions. For that, there are some standard principles defined by international organizations —as the Organisation for Economic Co-operation and Development (OECD) and the International Labour Organization (ILO)— to classify the main activity of people. These "ILO Principles" consider three aspects.

First, within the total population in a country, exists a working-age population, considered to be able and likely to work, commonly aged between 15 and 64 (though ages can change). People who are not counted are considered out of the labor market.

Second, the labor force comprises all working-age people, during a specified brief period of one week or one day (often known as "the reference period"), fulfill the requirements to be employed or unemployed. The labor force framework classifies the in-scope population into two mutually exclusive categories: "employed", "unemployed", and "inactive" or "out of the labor force". People are classified as "employed" if they did any work at all for pay or profit during a determined period of time (usually one hour, called *the one-hour criterion*) in the reference week. This includes paid work, self-employed work, or people who are temporarily not working but who have formally paid work (*e.g.* they have a salary, women who are on maternity leave, etc.). People are classified as "unemployed" if they were without either paid nor self-employed job, but they are willing to work, that is if they have been actively seeking job over the past four weeks, or they have a new job to start within the following four weeks. "Actively seeking for a job" means that an individual must use job-search methods other than looking at job advertisements.

The one-hour criterion in the definition of employment that attempts to cover all types of employment that may exist in a given country, particularly any kind of irregular work (Hussmanns, Mehran, & Varmā, 1990). It is also known as "The ILO Principle" because ILO promotes it as part of the XIII International Conference of Labor Statisticians (ICLS) since 1982. This criterion is also fundamental in the definition of unemployment as a situation of total lack of work. In the labor force status framework, the definitions of employment and unemployment are interrelated. Likewise, the one-hour criterion entails the reference period to which this criterion should be applied, which could be either one week or one day.

Third, the labor force comprises employed and unemployed, which define the supply of labor in a market. Anyone who is not classified as a member of the labor force (employed or unemployed) is counted as "out of labor force", or inactive. For instance, many who are not in the labor force are going to school or are retired. Another people keep out of the labor force because of family responsibilities (ILO, 2004).

Usually, each country has its own mechanism for measuring labor market conditions trying to follow international standards and guidelines. However, there are several cases of countries with problems in their long-run national statistics. Problems are consequences of redesign methods to measure labor force status and imply data comparability problems.

New Zealand is a country that had introduced changes in its methods to study labor market and officially admit its problems to analyze data in long-run terms. It introduced an official labor force measurement in 1985 with the Household Labour Force Survey (HLFS) (Statistics NZ, 2016a, 2016b). In 2016, HLFS was redesigned, the first substantial change since it was introduced.

A significant change arising from the redevelopment of the HLFS comes as a result of improved accuracy in identifying active job seekers, but this has implications for unemployment and labour force participation statistics. HLFS now classifies looking at job advertisements on the Internet as not actively seeking work[1], and so, Statistics New Zealand (department charged with the collection of statistics) has removed them from the unemployment rates. Previously, responses that specified using Internet in job seeking were captured in an 'other' category and, consequently, classified as 'actively seeking'. Another important change corresponds to the reasons considered for not currently being employed, *i.e.* to be unemployed or out of the labor force. They were updated questions and answer categories to collect more detail on respondents' reasons for not participating in the labor force and their main activities.

Nigeria has also changed its method to measure labor force status. The Household Labour Force Survey (HLFS), survey that measures the labor market conditions, were conducted invariably from 1966 to 2015 under the responsibility of National Bureau of Statistics (NBS). During that period, to be considered employed, a person must meet the criteria of working at least 40 hours in the reference week.

However, in 2015 NBS introduced a revised method for computing labor statistics. These changes were two. First, NBS adopted 20 hours a week benchmark, as against the 40 hours that had been used historically. Second, the NBS also introduced a new measure calculating underemployment (Kale & Doguwa, 2015). Nigerian labor force data was previously presented in binary terms: the number of employed people versus the number of unemployed. With the revision, people working between 20 to 39 hours a week are listed as underemployed. As a result, unemployment rate is higher and not comparable with rate measured whether it would be measured as ILO guidance. As consequence, the *ILO principle* puts Nigeria in a difficult situation, as most of its citizens, who put in more than the required benchmark hours in menial jobs, do not earn enough to be considered employed (NBS, 2014, 2015).

Problems using labor long-data series are also present in some Latin American statistics. Important cases to consider are Mexico, Colombia, and Brazil.

[1] It was jointed with 'Looked at job advertisements in newspapers' in 'Looked at job advertisements'

In Mexico, statistical difficulties arise in a lack of long-term data series: since 1972, the country has had five different surveys that have attempted to measure labor market conditions. Results of these surveys are incomparable, especially before 1988, due to differences between sampled target populations, data collection periods, and questionnaire designs[2]. After four different surveys and discontinuities due to the redesign, in 2005 the National Bureau of Statistics (Instituto Nacional de Estadística, Geografía e Informática, INEGI) introduced the National Survey of Occupation and Employment (Encuesta Nacional de Ocupación y Empleo, ENOE), which combines ENEU and ENE, being the current instrument used for labor force measurement.

It is also important to remark that Mexico has had one of the lowest unemployment rates in the world, something that intuitively does not correspond to the level of development nor with the very low rate of Mexican economic growth over the past three decades (Fleck & Sorrentino, 1994; Gutiérrez, 2009; Martin, 2000; Márquez, 2015). This explains why unemployment indicators have never gained much acceptance in Mexico.

Colombia, as Mexico, has redesigned its labor statistics since the mid-1970s. Two are the main data problems: official statistics produced by national statistical institutes are not comparable over time, in addition to the fact the changes were implemente even within the period when surveys have been collected. Specifically, since the mid-1970s, three surveys have been conducted and the three of them have had changes in measurement of labor force status in early 2000s[3].

In 2006, it was introduced the Great Integrated Household Survey (Gran Encuesta Integrada de Hogares, GEIH), which replaced the three main existing household surveys (ECH, the Encuesta de Ingresos y Gastos, and the Encuesta de Calidad de Vida) (DANE, 2009; Salazar, 2006). Also, urban areas increased from 13 to 24 (13 areas included quarterly and 11 twice a year).

The redesign of survey occurred in 2000 with ECH, but specific changes in coverage and representativeness only occurred with GEIH (DANE, 2012). Most important changes are two. First, period of data collection went from quarterly to monthly. This change allowed increasing the frequency of data production, facilitating the follow-up of short-run tendencies. Second, the measurement of employment and unemployment adopted ILO

[2]In 1972, Mexico began to measure unemployment with the National Household Survey (Encuesta Nacional de Hogares, ENH). A year after, in 1973, is replaced by the Continuous Workforce Survey (Encuesta Continua sobre Mano de Obra, ECMO). And, in 1974 it was once again replaced by the Continuous Occupational Survey (Encuesta Continua sobre Ocupación, ECSO), that covered only the three larger cities (Mexico City, Guadalajara, and Monterrey). In 1983, it was introduced the National Survey on Urban Employment (Encuesta Nacional de Empleo Urbano, ENEU), which at the beginning considered only 12 larger urban areas to reach 45 cities in 2000. Since 1995, also, survey has been conducted annually.

[3]Labor market conditions in Colombia were measured since 1976 using the National Household Survey (Encuesta Nacional de Hogares, ENH), including Santa Fe de Bogotá, Cali, Medellín and Barranquilla. In 1978, it included rural areas but just in 1988 this was repeated, and in 1980s urban areas rose to 13. This survey was redesigned and, later, replaced in 2000 by the Continuous Household Survey Continuous Household Survey (Encuesta Continua de Hogares, ECH). Its frequency passed from twice a year to continuous survey, and in 2001 geographic coverage rose from 13 larger urban areas to national coverage.

guidelines: thus, to be considered "employed" people must have worked for one hour on one day, at least, during the reference period. To be considered "unemployed", on the contrary, is not merely not having a job nor being actively seeking for a job, it is also required to be available to work immediately. In addition, it measures *hidden unemployment*, which refers to people who are available to start to work but have not being actively seeking for a job.

For the case of Brazil there is also a lack of a long-data series. Unlike Mexico, Brazil has two periods of comparable data which are not comparable to each other. In 2016 began a new series, not comparable with the earlier ones[4].

After two prior methods of measurement of labor force status —the PNAD and the PME—, in 2016 the Brazilian National Bureau of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística, IBGE) introduced the Continuous National Household Sample Survey, known as the Continuous PNAD (Pesquisa Nacional por Amostra de Domicílios Contínua, PNAD Contínua), as the unique data source for addressing labor force status.

In this case, the redesign of questionnaires and surveys (and their measurements) implied changes in data collection periods (reduce their regularity), in the geographical coverage (it increased it), and in the final statistics for labor force status (IBGE, 2015a). If the PME considered only six metropolitan areas with monthly frequency, Continuos PNAD considered 3.500 cities quarterly. Also, conceptual changes affected national statistics. First, Working-Age Population was redefined. Until 2002, the PME considered people aged 15 years and older, later PME and PNAD considered population aged 10 years and older. However, Continuous PNAD rose this limit to 14 years of age. Additionally, to be counted as an "unemployed" also varies across surveys. PME considered as "unemployed" to someone who is available to work immediately and looking for a job in a 30-day period, while Continuous PNAD does not consider period to seeking for a job (IBGE, 2015b).

Aforementioned examples summarize two important assumptions that back up this work. On the one hand, they illustrate the importance of keeping updated labor market data sources and their role in how data affects labor market conditions. On the other hand, they enlighten the trade-offs between the redesign of measurement methods and the production of comparable data to obtain a long-run perspective.

[4]In 1967 was introduced the National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios, PNAD), the first household survey collected in Brazil which considers questions about labor market. Survey was conducted irregularly next decades, and in 1980 it started the Monthly Labor Survey (Pesquisa Mensal de Emprego, PME) as labor force survey, which considered six metropolitan areas (Belo Horizonte, Porto Alegre, Recife, Rio de Janeiro, Salvador, and São Paulo). Its methodology was redesigned in 2002 following ILO guidelines, losing comparability over time.

1.2 Overview of the Labor Market in Chile

Labor market measurements are closely linked to characteristics of the region or the country where they are used. This is why it is important to consider labor market conditions in Chile in recent years to provide a framework for our subsequent analyses.

First, we have to consider that the Chilean labor market is the most flexible in Latin America. This conclusion is a relatively new discussion who rase up after several investigations conducted during the last decade or so. On this, it is remarkable the results work of the World Bank's Rigidity of Employment Index[5] (Sehnbruch, 2006).

Second, downwards in the labor market dynamics are a consequence of the overall performance of the national economy (due to cyclical component), although it does not mean it downplays the role of institutions. Considering different periods, authors such as Bravo et al. (2005); Cowan et al. (2005); Landerretche (2007); Naudon and García (2012); Sehnbruch (2006) have pointed out this issue.

Third, an important aspect of any labor market is the unemployment rate. Labor market responds to economic fluctuations through more variables than to the unemployment rate solely; however Chilean labor market adjusts to recent economic crises predominantly through the unemployment rate, as opposed to wage levels (Cowan et al., 2005; Sehnbruch, 2006). Cowan, Micco, Mizala, Pagés, and Romaguera (2005) analyzed labor dynamics from 1998 to 2002, focusing on the causes of the increase in unemployment rate, the most severe and persistent effect of the crisis. As a conclusion, the authors defined the fall in job creation as the main cause of this phenomenon, which did not match with a similar fall in the Economically Active Population, not finding any evidence related to structural changes in job creation. Nevertheless, they did find evidence of the influence of the rise in the minimum wage between 1997 and 2000 (the share of workers affected is nearly double for inexperienced and uneducated workers). Covering similar period, 1996-2004, Bravo, Ferrada, and Landerretche (2005) studied the relation between labor market and economic cycles. They found that unemployment was strongly closer with job-to-job transition rates. Namely, it increased probability of being unemployed while it declined overall job-to-job transition rate. Also, the increase in the unemployment after the Asian Crisis was mainly due to the job finding rate —and not by an increase in the job separation rate—, following Hall (2005). Landerretche (2007), on the other side, studied job creation and destruction since the demand side covering the same period as Bravo et al. (2005), and controlling by firm size. He concluded that smaller enterprises have a much higher probability of destroying jobs and a lower probability of creating them (weaker in urban industries such as construction, manufacturing, commerce and services, the main areas in the Chilean economy), and job destruction is more likely due to bankruptcy.

[5]The Rigidity of Employment Index measures the regulation of employment, specifically the hiring and firing of workers and the rigidity of working hours. This index is the average of three subindexes: a difficulty of hiring index, a rigidity of hours index, and a difficulty of firing index. The index ranges from 0 to 100, with higher values indicating more rigid regulations. Chile scored 18 points in 2009, as same as Uruguay, and merely under Colombia, with 10 points. These countries have by far the lowest scores in Latin America, but still higher than most Anglo-Saxon countries.

Sehnbruch (2006) analyzed the development of unemployment during 1990-2003, distinguishing between the situations of the different members of the household. She found that the unemployment rate increased much more dramatically among household heads (or primary earners); the family members least affected by the rise of unemployment are the children of the household heads. Also, the unemployment rates of spouses increased more than the average rate, but less than it did among household heads. During 1993 to 2009, Naudon and García (2012) observed that the probability of losing a job increases and the probability of finding one decreases during recessions, but the labor force is not particularly sensitive to stagnations.

These results are relevant to study labor market, however, they just consider data until 2009 (collected with discontinued method, detailed in chapter 2). Results, therefore, do not reflect completely the Subprime Mortgage Crisis and its effects. After the change in the measurement of the labor force status in 2010, results have been under discussion due to the differences between both methods, which do not allow comparisons directly. A comparison method has so far been proposed by OECD (2017), which compares data for the period 2007-2016 using a weighted average and series adjusted for the break between 2009 and 2010 using new-to-old splicing factor available for Q4 2009 for employment, unemployment and population series by five-year age groups and gender alike. Nevertheless, employment and unemployment rates have a marked shift across 2009 and 2010 (period of break).

After 2010, the unemployment rate has been low despite the poor economic growth performance since 2014. Marcel and Naudon (2016) examine this apparent contradiction, and considering discontinuity in data, studied the evolution of labor market in two stages: between 1996 and 2009, and between 2010 and 2016. Considering cyclical aspects (*i.e.* the manner that economic activity modifies competitive equilibrium) they observed that changes in unemployment rate are a reflection of the sharp changes in the labor market only if job destruction increased sharply. Authors attributed this downward trend in unemployment rate to the growing importance of those older than 55 years of age in the labor force (combined with a declining importance of those younger than 25 of age). And so, findings suggest that differences in the evolution of unemployment rate and the economic growth are lower than they may appear at the first sight.

Aforementioned results are key elements to support the proposal of this thesis to compare both measurements of labor states, in particular the approach made by Marcel and Naudon (2016), because it considers micro and macroeconometric elements and, at the same time, they propose a method to deal with data discontinuity.

1.3 Labor Market Dynamics

Behind macroeconomic labor dynamics, there are several population dynamics. As in most countries, unemployment in Chile is far from being equally distributed across the labor force. Some groups are far more affected than others. An important assumption made by Marcel and Naudon (2016), also important to this proposal, was that unemployment rate evolution depends on three components. Besides cyclical components, the other two components are how different social groups insert into the labor market and demographic issues and their later effect on aggregate indicators.

Because this proposal aims to deal with the differences in the measurement of the labor force status —which is based on the probability of being in any of the three labor force statuses— this section presents a review of the individual and household characteristics that affect it.

First, individual characteristics correspond to demographic profiles. These characteristics illustrate individual conditions, connected to differences in the labor market insertion. As Cowan et al. (2005); Sehnbruch (2006) have found, the insertion in Chilean labor market differs by gender, which may be attributable to men and women's social roles (among others). Naudon and García (2012) further explain that women move in and out of the labor force more frequently because they appear to be more willing to accept less favorable working conditions when offered. Also, some authors as Angrist and Evans (1998); Heckman and Macurdy (1980) have presented evidence about negative causal effect of family size (fertility) on female labor supply. Cruces and Galiani (2005) report similar results on the effects of the number of children in the female labor supply by Mexico and Argentina. All these works have found that educational attainment (considered as educational stages) is lower among married women than among single ones. So, both gender and marital status are important to understand female labor supply.

An unstable insertion in the labor market is more closely related to low-skill levels. A worker's knowledge and skills —which come from his/her education and training— generate a certain stock of productive capital. The value of this productive capital is derived from how much these skills can valued and compensated in the labor market. Also, an unequilibrium between supply and demand of a particular training or skill level in several markets causes imbalances (Ehrenberg & Smith, 2009). Imbalances are closer to schooling since education credentials signal abilities, therefore, low-skilled people have problems to find a job in markets with high-skilled jobs. For example, since women have substantial interruptions of their careers, using proxies for experience such as the estimated time since one left school (potential experience) will understate gender differences in the labor market qualifications (Blau & Kahn, 2013). Landerretche (2007) found that women are not more fragile workers but they are discriminated during job creation periods; and schooling has a significant effect on reduce fragility.

In Chile, the unemployment rate during the Asian Crisis (and later) increased amongst less experienced workers, more skilled workers (considering Secondary and Higher Education), and among young people (principally in age group of 18-25) due to lack of experience (Sehnbruch, 2006). In fact, the unemployment rate experienced by high-skilled workers and low-skilled workers explain 41 percent of the increase in the overall unemployment rate during this period (Cowan et al., 2005), meanwhile young people are more likely to loss a job than adults (Naudon & García, 2012). So, whereas young people have no working experience, people aged 55 and over have gained too much experience for the labor market.

Young people, older workers, and women can be labeled as "secondary labor force groups" or as a part of a "secondary labor market", because these groups are are less closer to the labor market than men in age group between 25 and 54 years of age because their labor supply is more elastic compared to middle-aged males —sometimes called "primary labor market"—. When the secondary labor market becomes into an important share of the overall labor market as in Chile, it opens the door to a more "atypical" performance of the aggregate unemployment rate (Marcel & Naudon, 2016).

People decide how much labor they are willing to supply to the labor market (where and whether to work) and how much leisure they want for themselves. These decisions are related to household conditions, obviously. Remember, matters such as the number of children in a household affects negatively the female labor supply because women are the ones who principally do most of the household work and child care (Ehrenberg & Smith, 2009). In male labor supply, number of children (and their age) may affect positively their supply, because in gender roles men are providers and protectors (Boeri & Van Ours, 2008; Sehnbruch, 2006).

Another important element is the market conditions. The demand for labor in the market must deal with employers whose decisions about hiring are influenced by conditions in three markets: the labor, the capital and the product markets.

Macroeconomic factors may be behind differences in the efficient allocation of workers in the labor market, across states or regions within a country. For instance, Blanchard and Katz (1992) concluded that analyzing the growth rates of employment across US states, where some states have consistently grown faster than the national average, while some have grown more slowly. Booms and slumps for states are best described as transitory accelerations or slowdowns of employment growth. Growth eventually returns to normal, but the path of employment is permanently affected. These transitory changes in growth lead to transitory fluctuations in relative unemployment and wages. They also found that an increase in the demand leads to an increase in employment, where labor force participation and unemployment rates remain constant. The dominant adjustment mechanism is labor mobility, rather than job creation or job migration.

To estimate the probability of being in a certain labor force status s, we include the term *aggregate effect*, taking the "peer effects" as reference. Peer effects refer to the causal effect of group characteristics on individual outcomes (Angrist & Pischke, 2009), and they have received intense attention in education research. In particular, Ammermueller and Pischke (2006), studied the connection between classmates' family backgrounds (measured by the average number of books in their homes) and student achievement in European primary schools. The average number of books is a feature of the home environment that predates test scores and is, therefore, unaffected by school-level random shocks, avoiding any identification problems — a typical matter when estimating peer effects regressions (Angrist & Pischke, 2009)—.

The role of social networks and referrals in the labor market has been studied in several contexts. First, job-finding processes through personal contacts of local interactions and peer effects in the space of social relations has been studied (Calvó-Armengol & Jackson, 2004). Second, and likewise, the role of physical space described by urban neighborhoods (Topa, 2011). Finally, literature also has addressed personnel economics, in work that examines employers' staffing practices and firms' internal labor markets as productivity or wages (Bandiera, Barankay, & Rasul, 2010; Mas & Moretti, 2009).

The group effects on the individual are not original, but they are original in their use to estimate the probability of being in certain labor force state. Because people do not directly affect each other in the labor market, at least in macroeconomic dynamics, peer effect notion is not appropriate. On this, let us define aggregate effects as an unobservable shock that affects people with the same characteristics j at certain time t. So, people being in group j have an equal chance of being in a certain labor force status s in certain period of time t. These effects particularly come from labor market itself, unlike other indicators of Chilean economic activity, as IMACEC and INACER. Also, indicators of economic activity indicate trend, whereas proposed models in this work consider seasonal variables. Consequently, it is not simple work with these two dimensions of economy in a same model. At the same time, there are no test of independence to determine IMACEC and INACER trend behavior, that is why they cannot work in a forecasting method (considering an Out-of-Sample model).

Chapter 2

A Discussion of Labor Market Data

This thesis follows an inductive reasoning, where data during 2003 to 2009 is used to forecast counterfactuals for the period between 2010 and 2013. Also, forecasting results are compared with stylized facts for the same period to illustrate the difference between measurement, and then compared with trends exposed in section 1.2. In order to do that, main data source used is the National Employment Survey.

How to handle data (and differences between the two methods of measurement of the labor force status) becomes an important matter when using inductive reasoning. For that purpose, the first section focuses on explaining characteristics of both data sources. The second section describes the most important differences between the two surveys. The third section presents an empirical comparison between both measures, taking 2009 as reference year.

2.1 Data description

The data source used corresponds to the National Employment Survey, NES (or Encuesta Nacional de Empleo, ENE in Spanish), survey that measures the situation of working-age population in the workforce in Chile, according to their demographic characteristics. It is conducted by the Chilean National Bureau of Statistics, NBS (Instituto Nacional de Estadísticas, INE), who collects data each month using households as information unit.

NES uses a stratified two-stage probability-based survey design and proportional allocation of plots to strata, based on geographical conditions: administrative division in regions, on the one hand; and rural-rest of urban area-urban relation, on the other[1]. This

[1]As part of the sample update based on demographic characteristics, the NES sample has been redesigned after each decennial census.

survey structure allows national and regional representativeness (INE, 2010). The entire sample is divided into three independent subsamples, each of which represents the total population. Each month NES is collected in one of those subsamples alternately in successive periods during 18 months. So, households residing at an address selected into the sample are interviewed one month, not interviewed in the following two months, and then interviewed again (each household is interviewed six times over an 18-month period). That mechanism of collecting data is known as household rotation panels. Also, households are finally selected by Simple Random Sampling. An adult household member provides information of all household members.

The NES, which had been conducted regularly since 1966, was modified in 2010, —henceforth referred as New National Employment Survey, NNES— following technical guidelines proposed by the OECD, ILO, and the U.S. Bureau of Labor Statistics. The main objective of these methodological changes was to introduce labor force measurement methods used in developed countries, and so, to make the Chilean data comparable to data from other countries —as Chile became a member of the OCDE in 2010—. As a result of implementing those changes, the NNES design includes updated unemployment and participation rates, as well as the distinction between *traditional* and *non-traditional* employment, which are not comparable with previous data (NES).

2.2 Differences in Survey Design

Methodological changes introduced in NNES have sown seeds of doubt in conclusions drawn from from the new data due to the impossibility of comparing the NNES with the NES. For instance, recent analysis of the Chilean labor market, described in chapter 1, have used an approach —in the case of OECD (2017) statistics— or have analyzed the NES and the NNES separately —method used by Marcel and Naudon (2016)—. In that sense, Chilean data attained international comparability, but lost long-term comparability. Officially it does not exist a method to compare both sources of data.

To construct a continuous series from 2003 to 2013 with both surveys, it is important to keep in mind methodological differences between the NES and the NNES. This thesis has only considered theoretical differences between them more closely related to labour force status definition. Principal differences between the NES and the NNES are explained in Table 2.1.

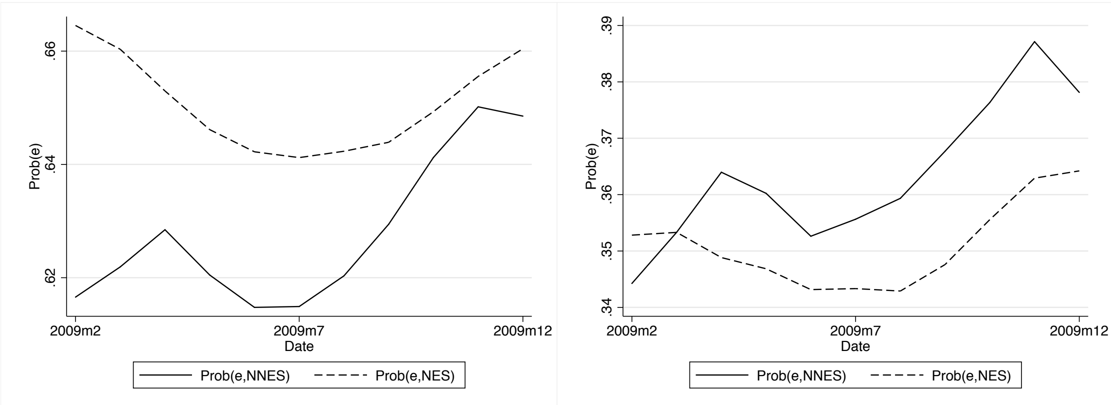
The key difference between both methods —and which makes impossible any kind of direct comparison— is how to determine labor force status. Thus, the NES considered people as employed if they did a work for pay or profit for the most important part of survey reference week, while the NNES, adopting the one-hour criterion, considers as employed people are those who have done any work at least one (paid) hour during the reference week. Because the condition to be considered employed is more demanding in the NES than the NNES, the latter tends to underestimate the unemployment situation, although this effect is different by gender.

Dimension	NES	NNES
Mode to delimit labor force status	Direct. Self-assessment answer	Indirect. It must fulfill several conditions asked in questions
Definition of employed	Surveyed declare have worked (paid) the most important part of reference week	Surveyed worked at least one hour for pay or profit
Employed status	People are considered as employed only if they have a job	Differentiate between *traditional* and *non-traditional* emploment, depending on if respondents identify their activity as work
Definition of unemployed	People who are not working during reference week, who have been seeking a job in the prior 8 weeks (they must have worked before the unemployment period); or people who are seeking a job for the first time	People who are not working during reference week but are currently available for work, and who have been actively seeking a job within the last 4 weeks (including reference week); or people who are seeking a job for the first time
Definition of inactivity	People are considered out of labor market if they are not employed and are not seeking a job	People are considered out the labor force if they are not employed and are not seeking a job. In addition, into inactivity condition, there are more detailed classification people according to reasons to inactivity and their willingness to work
Classification of Economic Activities	9 productive activity categories	17 productive activity categories. Introducing the use of ISIC (International Standard Industrial Classification of all Economic Activities), which allows international comparability
Compute Working Hours	Solely considers declared working hours by people	Computes actual working hours from usual working hours and changes in usual work hours
Sampling Frame	Defined by Population and Housing Census, collected every 10 years	Different for urban and rural area. To rural area keeps NES sampling frame based in Census. To urban area, frame is generated from Census and digital mapping (Mapcity and Dmapas, and municipal information about building permits). Digital mapping allows constantly updating

Table 2.1: Comparison differences between NES and NNES

2.3 Observed differences

An approach to illustrate differences in results between both surveys is analyzing data collected in 2009. That year, NBS collected both the NES and the NNES simultaneously. Although those surveys were carried out at the same time only during 2009, it is an appropriate framework to observe empirical differences between them and their influence in statistics evolution.



(a) Proportion(e) of men (b) Proportion(e) of women

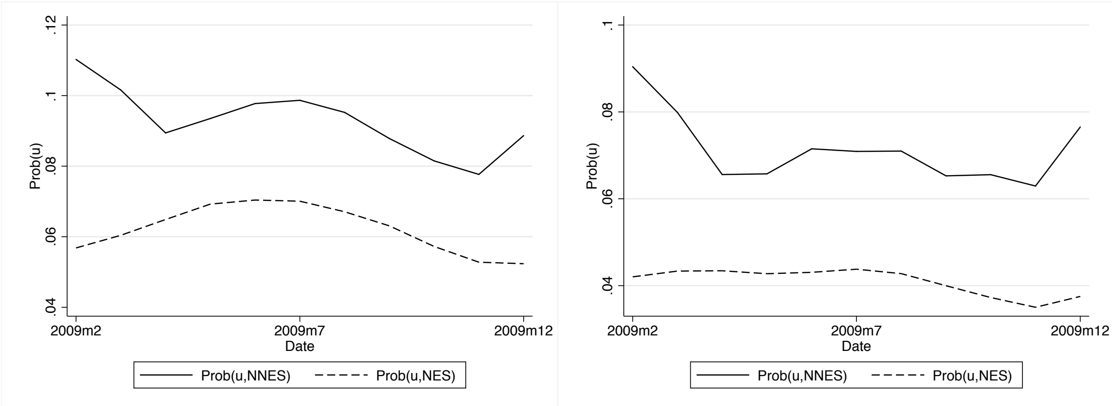
Figure 2.1: Proportion of Working-Age Population Employed 2009, NES and NNES

Tools used to measure social performance have a significant impact on possible explanation of women's labor insertion. As explained in section 1.3, women have a weaker integration into the labor market, being more affected by job insecurity and working less hours than they would like. So, while the NES was collected, women are more probably to be considered as unemployed or out of the labor market. In fact, the increase in women's probability of being employed can be explained merely by the change of method —as illustrates Figure 2.1— the same as female labor force participation, comparing women's chance of being employed and unemployed increase versus the probability of being out of labor market falls . Also, male participation rate evolution presents a similar trend, but does not change as does female's due mainly to changes in the probability of being unemployed.

The probability of being unemployed, presented in Figure 2.2, confirms the increase in the labor force participation considering NNES results, for both women and men. Also, reducing the period considered for seeking a job (and, hence, to be counted as unemployed) has some impact in unemployment evolution. So, it is more likely to be considered as "unemployed" in the NNES.

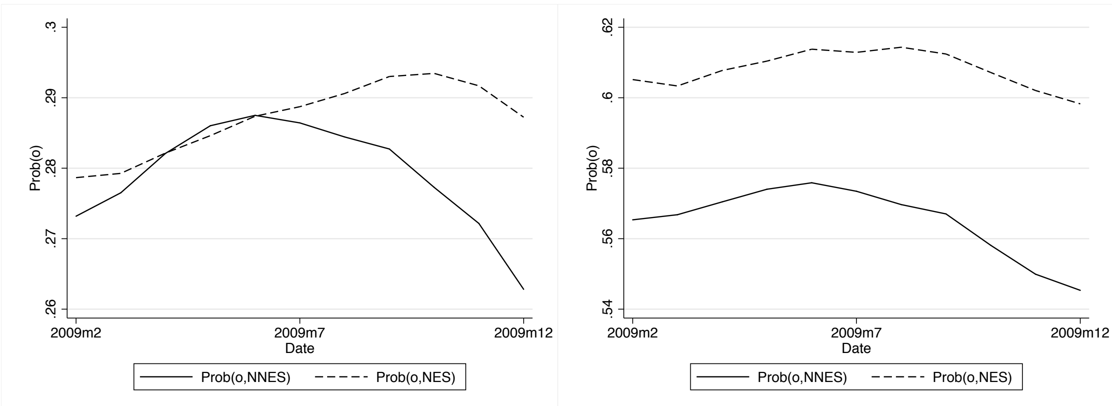
Considering differences between NES and NNES, models used to estimate counterfactuals of NES results in NNES period rely on three assumptions.

Assumption 1 (**Population**) *The NES and the NNES are drawn from the same target population with no structural change in its evolution —it has not been affected by shocks (if 2010 earthquake is considered as transitory shock)—.*



(a) Proportion(u) of men (b) Proportion(u) of women

Figure 2.2: Proportion of Working-Age Population Unemployed 2009, NES and NNES



(a) Proportion(o) of men (b) Proportion(o) of women

Figure 2.3: Proportion of Working-Age Population Out of market 2009, NES and NNES

Assumption 2 (**Comparability**) *The NES and the NNES use the same sample design, and since strata and sampling base weights design are based on same last census released, sample remains the same. Thus, changes, merely, correspond to those related to the measurement of labor force status.*

Assumption 3 (**Structure of Data**) *Although the NES and the NNES have panel data structure, models assume data as cross-sectional but used with natural temporal ordering (as time series structure).*

This thesis uses monthly frequency of NES and NNES, and analyses cover the period between 2003 and 2013. From 2003 because that year began the labor market recovery after the Asian Crisis. This crisis caused structural changes in the market: a higher unemployment rate, an increase in the labor force participation rates, and a constant increase in the creation of dependent jobs (Cowan et al., 2005; Landerretche, 2007). Nevertheless, working under the Assumption 1 implies some problems when data is being collected long way off changes, in particular, it may induce biased results. Therefore, analyses cover only three years of the NNES series.

Chapter 3

Empirical Strategy

This chapter presents the methodological proposal to deal with data problems explained in chapter 2. So, the first section details the construction of a comparable labor force status variable throughout the analyzed period. The second section describes how counterfactual probabilities are estimated by forecasting as well as their evaluation.

3.1 Creating a Labor Force Status Variable

This section presents in detail the design of a comparable labor force status variable. For that, observed labor force status variable in the NES is a good starting point for modeling an initial discrete response model, which allows obtaining parameters of prediction of labor status as a dependent variable. In this sense, the methodological proposal considers merely three possible statuses: employment, unemployment, and inactive or out of the labor force[1].

3.1.1 Binary Response Model

It is important to remark that it does not exist a long-term and nationally representative labor force data in Chile. Our proposed solution is to construct a counterfactual using an approximate model of labor force status estimated through predicted probability of being in one of the three labor force statuses. Making a counterfactual prediction begins with a discrete response model to obtain parameters of these probabilities. The variable to be explained corresponds to labor force status y, being a random variable taking on a finite

[1]The NNES considers sub-status. So, there are traditional, non traditional, and absent employed. Also, there are unemployed and who are looking for their first job. And out of labor force considers family, study, stationary, temporary, permanently sick, and retirement reasons, not available for work, and starter.

number of outcomes (Wooldridge, 2010), which permit modeling a qualitative variable with techniques for discrete variables.

In a binary response model, interest lies in the parametrization of probability of the event p depending on a vector X and on a a vector of parameters β. Conditional probability is given by:

$$p(X) \equiv \mathbb{P}[y = 1|X] \equiv G(X\beta) \tag{3.1}$$

where $G(\cdot)$ is the cumulative distribution function and maps the index $X\beta$ into the response probability. In particular, restricts $p_\mathrm{i} \in (0, 1)$ in Equation 3.1.

Functional form for $G(\cdot)$ defines a particular regression model. Models used in Economics of the form of Equation 3.1 are two: probit and logit. The difference between logit and probit is likely to be small, and there are practical reasons for favoring logit or probit (*e.g.* mathematical convenience), but it is difficult to justify the choice of one distribution or another on theoretical grounds (Greene, 2011). In this case, it is necessary to consider the fact that the NES sample is large enough, so there are differences in the distribution: logistic distribution has fatter tails than probit. Figure 3.1 illustrates their differences:

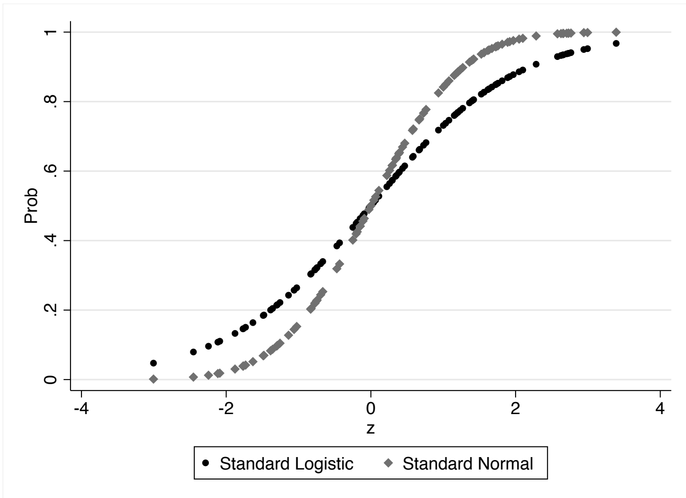


Figure 3.1: Comparison of logit and probit distribution

The density function of the standard normal distribution (cdf of probit model) has thinner tails than the density of the standard logistic probability density function. That makes larger marginal effects than logistic distribution. So, initial probabilities are estimated using probit model and $G(z) = \Phi(z)$.

Defined $G(\cdot)$, we can estimate the value of parameters β. Probit models usually estimate $\hat{\beta}$ using Maximum Likelihood because data distribution is well-defined by Bernoulli model, where exists a case of success $y = 1$ and a case of failure $y = 0$. To estimate $\hat{\beta}$ that maximize likelihood function[2], density function is:

$$g(y_\mathrm{i}|X_\mathrm{i}) = p_\mathrm{i}^{y_\mathrm{i}}(1 - p_\mathrm{i})^{1-y_\mathrm{i}} = \int_{-\infty}^{z} \phi(v)\mathrm{d}v = \int_{-\infty}^{z} (2\pi)^{-1/2} \exp(-z^2/2)\mathrm{d}z \equiv \Phi(z) \tag{3.2}$$

[2]From general Maximum Likelihood results, we know $\hat{\beta}$ is consistent and asymptotically normal.

From Equation 3.2, probabilities p_{i} and $(1-p_{\mathrm{i}})$ are $g(1) = p_{\mathrm{i}}^{1}(1-p_{\mathrm{i}})^{0} = p_{\mathrm{i}}$ and $g(0) = p_{\mathrm{i}}^{0}(1-p_{\mathrm{i}})^{1} = (1-p_{\mathrm{i}})$. Log-likelihood function is given by:

$$\mathrm{L}_N(\beta) = \sum_{\mathrm{i}=1}^{N} \left[y_{\mathrm{i}} \ln \Phi(X\beta) + (1-y_{\mathrm{i}}) \ln\left(1-\Phi(X\beta)\right) \right]$$

With this, $\hat{\beta}$ that solve First-Order Condition is:

$$\sum_{\mathrm{i}=1}^{N} \frac{y_{\mathrm{i}} - \Phi(X\hat{\beta})}{\Phi(X\hat{\beta})(1-\Phi(X\hat{\beta}))} \Phi'(X\hat{\beta})X = 0$$

Later, the measurement of marginal effects expects an instantaneous change in the dependent variable as a function of a change in a certain explanatory variable while keeping all the other covariates constant. Marginal effect in change of j-th regressor is dependent of an observation i, that is why it is necessary to specify how to obtain marginal effects. Marginal effects may be computed in means of vector X_{i}, or in average of marginal effects[3]. Second choice is used in this case, defined as:

$$\frac{\partial \bar{p}}{\partial \bar{x}_j} = \sum_{\mathrm{i}=1}^{N} \frac{N\phi(X\beta)}{N} \hat{\beta}_j$$

To predict labor force status $s = \{\mathrm{e}, u, o\}$, there are two models. The first model considers individual characteristics for individuals $\mathrm{i} = 1, ..., I$ —both personal and household characteristics—. This individual model defines the probability of individual i at time t to be in state s as:

$$\mathbb{P}[y_{\mathrm{i},t} = s] = \Phi(X\beta_{\mathrm{i},t}) \tag{3.3}$$

In this model, vector $X_{\mathrm{i},t}$ includes observed variables such as marital status, schooling, condition to be head of household, age, Household Dependency Ratio (HDR)[4], and two interaction terms: an interaction term between age and the HDR and the interaction term between schooling and the HDR.

The second model adds conditions of labor market through *aggregate effects*. Let us define the probability of being in status s according to model with aggregate effects as:

$$\mathbb{P}[y_{\mathrm{i},t} = s] = \Phi\left(\beta X_{\mathrm{i},t} + \lambda \bar{Z}^{s}_{a(\mathrm{i}),t} + \mu \bar{Z}^{s}_{r(\mathrm{i}),t}\right) \tag{3.4}$$

Besides $X_{\mathrm{i},t}$, there are two aggregate effects: $\bar{Z}^{s}_{a(\mathrm{i}),t}$ (aggregate effect to being in s by age) and $\bar{Z}^{s}_{r(\mathrm{i}),t}$ (aggregate effect to being in s by region of living). In Equation 3.4, λ and μ are vectors of parameters associated with aggregate effects.

[3]To estimate marginal effects, by Slutsky theorem, it is the same thing estimate on mean or individual mean, because $\operatorname{plim} g(x_n) = g(\operatorname{plim} x_n)$, but in small samples this equation is not necessary true, so individual marginal effects are computed and then, their mean.

[4]HDR is the ratio of dependents (people younger than 15 and older than 64) to the working-age people at household.

Definition 1 (**Aggregate Effects**) *Aggregate effect is the probability of an individual* i *in working-age population* $N_{\mathrm{i},t}$ *of being in a certain labor force status* s*, due to effect, or pressure, of group* $g(\mathrm{i})$ *in a certain time* t*. In other words:*

$$\bar{Z}^s_{g(\mathrm{i}),t} = \frac{\sum_{\mathrm{i}} \Phi^{-1}(P^s_{\mathrm{i},t})\mathbf{1}[g(\mathrm{i}) = g] - \sum_{j\neq \mathrm{i}} a_j}{\sum_{\mathrm{i}} \mathbf{1}[g(\mathrm{i}) = g] - 1}$$

There are two important considerations for empirical implementation. First, models have been estimated separately by sex, although using the same model. That is because the evidence presented in chapter 1 illustrates differences between men and women in their labor market insertion. Second, models have estimated equations that asymptotically reflect probabilities of being in each state with frequencies of people in labor state s in relation with the total working-age people according groups $g(\mathrm{i})$ such that $P^s_{\mathrm{i},t} \xrightarrow{p} \mathbb{P}[y_{\mathrm{i},t} = s]$. Also, to estimate s, each labor force status is considered as binary variable, where:

$$y_{\mathrm{i}} = \begin{cases} 1 & \text{if } y_{\mathrm{i}} = s, \text{ with probability } p_{\mathrm{i}} \\ 0 & \text{if otherwise, with probability } (1 - p_{\mathrm{i}}) \end{cases}$$

3.1.2 Generalized Response Model

Making counterfactuals with the prediction of individual probabilities is highly demanding, therefore, analyses consider probabilities at the macro level. In the case of a model with aggregate effects, this implies to generalize the model of Equation 3.4. Model assumes that $\bar{Z}_{g(\mathrm{i}),t}$ is captured by the average of a transformation of the probability of being in s according to a —age assumes values $a = 1, ..., A$— and r —region of living $r = 1, ..., R$—. Applying the inverse cdf to Equation 3.4 and aggregating individuals by group a, result is:

$$\begin{aligned}
\bar{Z}^s_{a,t} &= \frac{\sum_{\mathrm{i}} \Phi^{-1}(P^s_{\mathrm{i},t})\mathbf{1}[a(\mathrm{i}) = a] - \sum_{j\neq \mathrm{i}} a_j}{\sum_{\mathrm{i}} \mathbf{1}[a(\mathrm{i}) = a] - 1} \\
&= \frac{\sum_{\mathrm{i}} X_{\mathrm{i},t}\mathbf{1}[a(\mathrm{i}) = a]}{\sum_{\mathrm{i}} \mathbf{1}[a(\mathrm{i}) = a]}\beta + \lambda \bar{Z}^s_{a(\mathrm{i}),t} + \mu \frac{\sum_{\mathrm{i}} \bar{Z}^s_{r(\mathrm{i}),t}\mathbf{1}[a(\mathrm{i}) = a] - \sum_{j\neq \mathrm{i}} a_j}{\sum_{\mathrm{i}} \mathbf{1}[a(\mathrm{i}) = a] - 1} \\
&= \bar{X}_{a,t}\beta + \lambda \bar{Z}^s_{a,t} + \mu \sum_r \omega_{a,r} \bar{Z}^s_{r,t} \quad \text{with } \omega_{a,r} \equiv \frac{\sum_{\mathrm{i}} \mathbf{1}[a(\mathrm{i}) = a]\mathbf{1}[r(\mathrm{i}) = r]}{\sum_{\mathrm{i}} \mathbf{1}[a(\mathrm{i}) = a]}
\end{aligned}$$

The first term in the last line is the average of characteristics X for a group a, and the third term may be obtained from $\bar{Z}^s_{r(\mathrm{i}),t} \equiv \sum_r \bar{Z}^s_{r,t}\mathbf{1}[r(\mathrm{i}) = r]$. Then, $\bar{Z}^s_{r(\mathrm{i}),t}$ is substituted into the second line of equation, and so $\omega_{a,r}$ is the share of individuals of a in group r over all individuals of a. Whereas, aggregate effect by r is obtain by:

$$\bar{Z}^s_{r,t} = \bar{X}_{r,t}\beta + \lambda \sum_b \psi_{r,a} \bar{Z}^s_{a,t} + \mu \bar{Z}^s_{r,t} \quad \text{with } \psi_{r,a} \equiv \frac{\sum_{\mathrm{i}} \mathbb{1}[r(\mathrm{i}) = r]\mathbb{1}[a(\mathrm{i}) = a]}{\sum_{\mathrm{i}} \mathbb{1}[r(\mathrm{i}) = r]}$$

Then, aggregate effects for all a and r form a linear equation system:

$$\bar{Z}_t^s = \bar{X}_t \beta + \Gamma \bar{Z}_t^s$$

where

$$\bar{Z}_t^s \equiv \begin{pmatrix} \bar{Z}_{a=1,t}^s \\ \ldots \\ \bar{Z}_{a=A,t}^s \\ \bar{Z}_{r=1,t}^s \\ \ldots \\ \bar{Z}_{r=R,t}^s \end{pmatrix} \quad \bar{X}_t \equiv \begin{pmatrix} \bar{X}_{a=1,t} \\ \ldots \\ \bar{X}_{a=A,t} \\ \bar{X}_{r=1,t} \\ \ldots \\ \bar{X}_{r=R,t} \end{pmatrix} \text{ and } \Gamma \equiv \begin{pmatrix} \lambda I_A & \mu \Omega_{A \times R} \\ \lambda \Psi_{R \times A} & \mu I_R \end{pmatrix}$$

with

$$\Omega \equiv \begin{pmatrix} \omega_{1,1} & \ldots & \omega_{1,R} \\ \ldots & \ldots & \ldots \\ \omega_{A,1} & \ldots & \omega_{A,R} \end{pmatrix} \text{ and } \Psi \equiv \begin{pmatrix} \psi_{1,1} & \ldots & \psi_{1,A} \\ \ldots & \ldots & \ldots \\ \psi_{R,1} & \ldots & \psi_{R,A} \end{pmatrix}$$

Γ joins aggregate effects and their weights in population Ω and Ψ. These effects do not depend on i —and have not relation with $\bar{X}_t$—, but do depend on a group of individuals effects. Both in the models for men and women, Γ is a $6+6$ matrix with a 12×12 system of equations with possible values of a and r. The solution for aggregate groups is:

$$\bar{Z}_t^s = (I_{A+R} - \Gamma)^{-1} \bar{X}_t \beta$$

Then, an aggregate estimator is obtained by properly weighting the sectorial estimates. Hence, from model with aggregate effects, probability of being in labor force state s given group $g(\mathrm{i})$ is as follows:

$$\widehat{P}_{g(\mathrm{i}),t}^s = \Phi \left(\sum_{g=1} \widehat{\bar{Z}}_{g(\mathrm{i}),t}^s \right) \kappa_{g(\mathrm{i}),t} \quad \text{with} \quad \kappa_{g(\mathrm{i}),t} = \frac{\sum_{\mathrm{i}} \mathbf{1}[g(\mathrm{i}) = g, t]}{N_t} \tag{3.5}$$

3.2 Forecasting

To estimate the counterfactual prediction, and then, to assess models' performance *forecasting* is used. By forecasting, it is possible to make a prediction about the value of a random variable y at horizon of time h, conditional on the information set in period t. So, let $\{y_t\}$ denotes the series to be forecasted and let $\{y_{t+h|t}\}$ denotes the forecast of $y_{t+h|t}$:

$$y_{t+h|t} = f(\Omega_t)$$

A successful forecasting requires that there are regularities to be captured; these regularities have to be informative about the future; the proposed method captures these regularities, an yet it excludes non-regularities (Pindyck & Rubinfeld, 1998). Also, traditional theory of Economic Forecasting is based on two key assumptions. In the first place, *the econometric model is a good representation of the economy.* It must be noticed that we have implicitly assumed that econometric models —both individual and with aggregate effects— are good representations of the economy. Even so, we must test this strong assumption. Second, *the structure of the economy remains relatively constant.* This assumption is also strong, because a parametric econometric model is completely described by its parameters (B. Hansen, 1992). Likewise, it requires to be proved, revising whether the estimated coefficients are stable over time[5].

3.2.1 Model Fitting and Error Estimation

To estimate the parameters of a model to subsequently estimate counterfactuals, and later, to check whether the first assumption of forecasting is fulfilled, *i.e* if econometric model is a good representation of the economy, an In-Sample forecast is used.

In-Sample forecast considers the expected value of the random variable observed by the sample, and it is used to obtain the estimates of the parameters (Inoue & Kilian, 2005). An In-Sample because forecasting is done with data used to develop the model. So, $\widehat{P}_t^s$ is estimated using period 2003-2009 by month: the model with aggregate effects is obtained with $\widehat{\beta}$, $\widehat{\lambda}$, $\widehat{\mu}$, whereas the individual model is estimated with $\widehat{\beta}$.

An important part of the analysis of economic forecasts is to check whether the forecasting model still performs well in future Out-of-Sample periods. The first measure of a model accuracy is the Root Mean Squared Error (RMSE), which is the standard deviation of the residuals (prediction errors). Thus, the RMSE is always no-negative, and values closer to zero are better (P. R. Hansen & Timmermann, 2012). In this case, the RMSE compares proportion of declared labor force situation between 2003 and 2009[6] and results of models ($\widehat{P}_t^s$). By definition, the RMSE is expressed as:

$$\text{RMSE}_t^s = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{P}_t^s - P_t^s\right)^2}$$

The RMSE also helps to decide which labor force status is predicted with least precision. This labor force status is then fitted, computing it as the complement of two other probabilities that assume that the complement of an event E is the set of possible outcomes not in E. Thus:

$$\widehat{P}_{\mathrm{i}}^s = 1 - \widehat{P}_j^s - \widehat{P}_k^s$$

Another metric to test accuracy is the Theil inequality coefficient, also known as Theil's U statistic. Based on the RMSE, it provides a descriptive measure of forecasting accuracy with values ranging from zero to one. If $U = 1$, indicates the worst possible degree of

[5]Instability may reflect structural phenomena (model misspecification, omitted variables, measurement error, etc.) or punctual events (oil crisis, economic policy measures, new regulations, etc.).

[6]Observed proportion of being in any labor force state is given by $P^{s=\mathrm{i}} = \frac{N_{s=j}}{N_{s=\mathrm{i}}+N_{s=j}+N_{s=k}}$

forecast inaccuracy, while $U = 0$ there are no forecast errors. The advantage of the Theil's U is that it is unit less as it compares to the RMSE. The equation for the Theil's U is:

$$\mathrm{U}^s = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\widehat{P}_t^s - P_t^s\right)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\widehat{P}_t^s)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(P_t^s)^2}}$$

Theil's U statistic can be decomposed into three *proportions of inequality*. Let us define them as:

$$\mathrm{U}_M^s = \frac{(\bar{\widehat{P}}_t^s - \bar{P}_t^s)^2}{\frac{1}{T}\sum_{t=1}^{T}(\widehat{P}_t^s - P_t^s)^2} \qquad \text{Bias proportion}$$

$$\mathrm{U}_S^s = \frac{(\widehat{\sigma}_t^s - \sigma_t^s)^2}{\frac{1}{T}\sum_{t=1}^{T}(\widehat{P}_t^s - P_t^s)^2} \qquad \text{Variance proportion}$$

$$\mathrm{U}_C^s = \frac{2(1-\rho)\widehat{\sigma}_t^s \sigma_t^s}{\frac{1}{T}\sum_{t=1}^{T}(\widehat{P}_t^s - P_t^s)^2} \qquad \text{Covariance proportion}$$

Proportions of inequality are useful to uncover the potential sources of predictive errors, and range from zero to one and always sum one. U_M indicates the degree of systematic error. The second component, U_S, measures the ability of the forecast to replicate the degree of variability in the series of interest. Last, U_C gauges the degree of unsystematic error within the various forecasts. A good set of forecasts will have almost no bias proportion and little or no variance proportion, and so, any remaining error will be due to unsystematic variations in the data (Pindyck & Rubinfeld, 1998).

(a) Prob$(s, X\beta)$ for women

Prob	**RMSE**	**U^s**	**U^s Proportions**		
			U$_M$	**U$_S$**	**U$_C$**
$s =$ e	0.0083	0.0127	0.2346	0.00008	0.7653
$s = u$	0.0041	0.0588	0.0056	0.00011	0.9943
$s = o$	0.0101	0.0079	0.0267	0.00011	0.9731

(b) Prob$(s, X\beta)$ for men

Prob	**RMSE**	**U^s**	**U^s Proportions**		
			U$_M$	**U$_S$**	**U$_C$**
$s =$ e	0.0071	0.0054	0.1904	0.00006	0.8096
$s = u$	0.0050	0.0584	0.0103	0.00014	0.9895
$s = o$	0.0053	0.0094	0.0733	0.00003	0.9266

Table 3.1: Forecast Accuracy to Prob$(s, X\beta)$ 2003-2009

A model which produces small values of the forecast evaluation statistics —both RMSE and Theil's U— is judged to be a good model. Table 3.1 and Table 3.2 present accuracy measures to individual model and model with aggregate effects, respectively.

If $U = 0$, $\widehat{P}_t^s = P_t^s$ for all t and its adjust is perfect. To any value of $U > 0$, ideal distribution to inequality statistic is $U^M = U^S = 0$ and $U^C = 1$ (Pindyck & Rubinfeld, 1998). An individual model has better adjust than a model with aggregate effects both to men and women, comparing RMSE and Theil's U. Decomposing Theil's U into proportions of inequality, $\widehat{P}_t^s$ estimated by individual model have a higher proportion of covariance, and due to measures unsystematic error, is regarded as a signal for a better model. On its side, $\widehat{P}_t^s$ was estimated using model with aggregate effects have problems with bias proportion, which indicates problems with systematic error; and have a moderately problem with the variance proportion, which indicates that the actual value of the variable has fluctuated considerably, not good indication.

(a) Prob(s, Z) for women

Prob	**RMSE**	**RMSE$^{\mathbf{a}}$**	**U$^{\mathbf{s}}$**	**U$^{\mathbf{s,a}}$**	**U$^{\mathbf{s}}$ Proportions**		
					U$_{\mathbf{M}}$	**U$_{\mathbf{S}}$**	**U$_{\mathbf{C}}$**
$s =$ e	0.0915	0.0054	0.1606	0.0082	0.6060	0.0519	0.3421
$s = u$	0.0118	0.0040	0.1689	0.0577	0.0266	0.8594	0.1140
$s = o$	0.0961	0.0052	0.0702	0.0041	0.4783	0.3952	0.1265

(b) Prob(s, Z) for men

Prob	**RMSE**	**RMSE$^{\mathbf{a}}$**	**U$^{\mathbf{s}}$**	**U$^{\mathbf{s,a}}$**	**U$^{\mathbf{s}}$ Proportions**		
					U$_{\mathbf{M}}$	**U$_{\mathbf{S}}$**	**U$_{\mathbf{C}}$**
$s =$ e	0.2267	0.0080	0.1454	0.0060	0.6100	0.1200	0.2700
$s = u$	0.0153	0.0067	0.1334	0.0645	0.0741	0.7091	0.2168
$s = o$	0.2351	0.0057	0.7112	0.0101	0.7632	0.1883	0.0486

[a] Adjusted prediction errors

Table 3.2: Forecast Accuracy to Prob(s, Z) 2003-2009

To solve problems detected with U_M of $\widehat{P}^s_{g(\mathrm{i}),t}$ estimated by models with aggregate effects, prediction obtained with Equation 3.5 is adjusted by parameters estimated in an Ordinary Least Square model. Using OLS model assures adjusted predictions at the means, and at the same time, is able to capture general dynamics of probabilities. So, final prediction is as follows:

$$\widehat{P}_t^s = \delta_0 + \delta_1 \widehat{P}_{a,t}^s + \delta_2 \widehat{P}_{r,t}^s \tag{3.6}$$

where δ is estimated in same period to secure fitting of model inside the NES. Accuracy measures to this correction are RMSE^a and U^a in Table 3.2, and they reveal an improvement in accuracy of predictions.

Considering all measures of accuracy —and adjusted predictions from model with aggregate effects—, let us define predicted status with least precision to be computed as complement probability. To estimates from individual model, complement probability to women is to be out of labor market, and to men is to be employed. Whereas modeling with aggregate effects, complement labor force status —both for men and women— is to be unemployed.

3.2.2 Stability

To test the second key assumption of forecasting, *i.e.* the structure of the economy is constant, and Assumption 1 —closely related—, is also used an In-Sample forecasting. It tests whether $\widehat{\beta}$, $\widehat{\mu}$, and $\widehat{\lambda}$ are stable over a period if there are no structural breaks (permanent change in the parameter vector of a model). Recursive techniques for a linear regression with k regressors start from:

$$y_t = x_t'\beta_t + \varepsilon_t$$

where β is estimated using i $= k + 1$ to T observations, and k represents the number of coefficient of the regression equation.

A large body of literature has emerged in relation to the develop of tests of model stability. Two important tests are the CUSUM and the CUSUM-of-Squares, based on recursive residuals.

The cumulative sum CUSUM of the recursive residuals assumes under the null hypothesis $\beta_t = \beta_0$ with $t = 1, ..., T$. The recursive residuals are standardized i.i.d. with mean of zero and constant variance proportional to $t - k - 1$, *i.e.* there are no structural breaks. Then, test takes the cumulative sum and plots its value against the upper $[k, \pm 3a\sqrt{(t-k)}]$ and lower $[k, \pm a\sqrt{(t-k)}]$ bounds with $a = 0.948$ for $\alpha = 0.05$, or 95% confidence interval, at each point (Perron, 2006).

To detect non-random deviations, those that do not necessarily come from a structural change in coefficients, the CUSUM-of-Squares (CUSUMSQ) test exists. This test is an alternative measure, although not equivalent to using CUSUM. The series of CUSUMSQ has an expected value that goes from zero at $t = 1$ to the end of the sample, $t = T$, and behaves like a $\chi^2(t)$. The expected value of S_t under the null hypothesis is $\mathbb{E}[S_t] = (t-k)/(T-k)$, which goes to zero at $t = k$. The significance of departures from the expected value line is assessed by reference to a pair of lines drown parallel to the $\mathbb{E}[S_t]$ line at a distance a above and below. This value depends on both the sample size $T - k$ and the significance level α.

In order, let us define the forecast variance of recursive residual estimated $\varepsilon_t = y_t - x_t'b_{t-1}$ as:

$$\sigma^2_{\mathrm{e}_t} = \sigma^2\sqrt{1 + x_t'(x_{t-1}'x_{t-1})^{-1}x_t}$$

and let us compute recursively random variable w_t, which is the t^{th} standardized recursive residual:

$$w_t = \frac{\varepsilon_t}{\sqrt{1 + x_t'(x_{t-1}'x_{t-1})^{-1}x_t}}$$

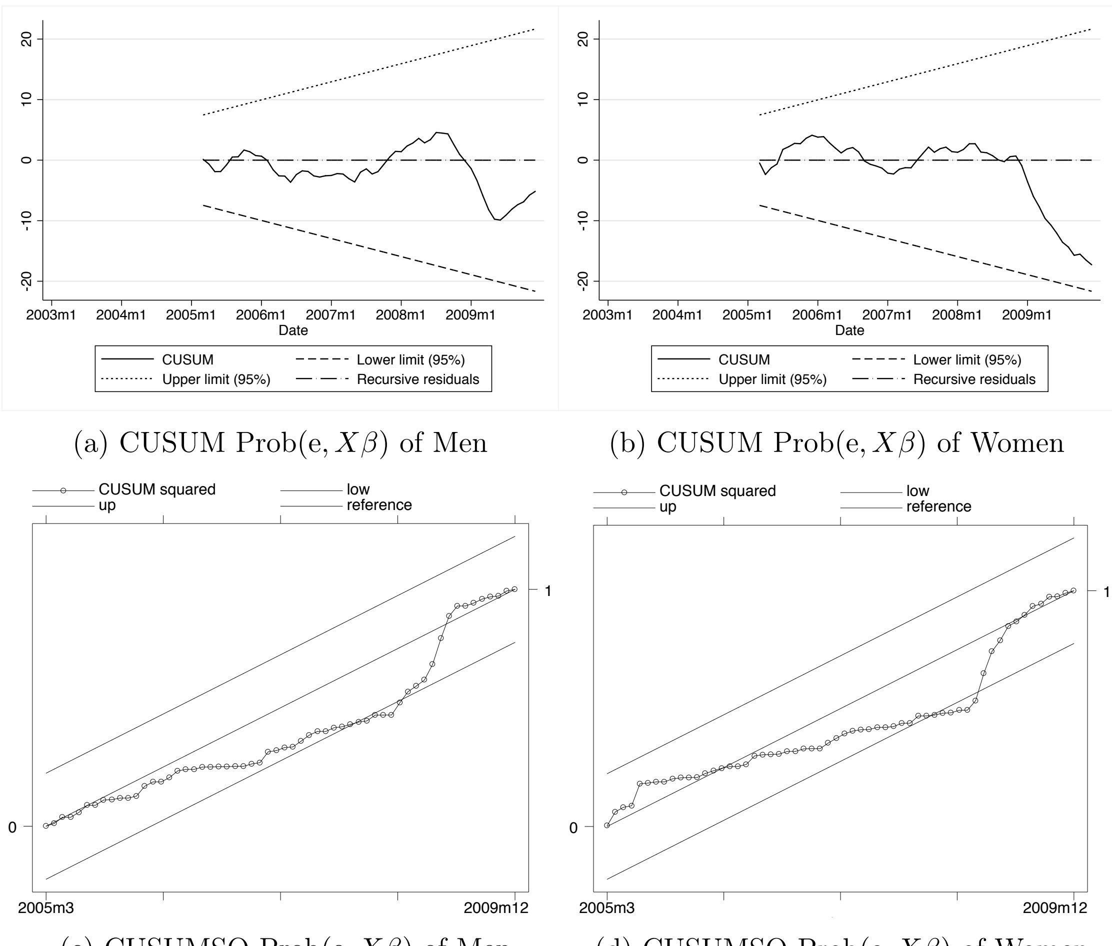
So, CUSUM test compute:

$$W_t = \sum_{j=k+1}^{T} \frac{w_t}{\widehat{\sigma}}, \quad \text{where } \widehat{\sigma} \text{ is the estimated variance of } w_t.$$

CUSUMSQ is based on cumulative sums of squared residuals:

$$S_t = \frac{\sum_{k+1}^{t} w_j^2}{\sum_{k+1}^{T} w_j^2} \qquad \text{with} \quad t = k+1, ..., T$$

Under the Assumption 3, it is assumed that the structure of data is not exactly a time-series structure. In order to estimate tests of model stability, data needs to be a time-series process. To create a series of data points indexed in time order, we use intuition behind the Taylor Series. So, estimation of tests uses instead some other function $F(x)$ that is both simple and a good approximation to $f(x)$ for x close to expansion point x_0. The Taylor series for a function is useful because we need to approximate the value of the function near x_0: in this case, the approximating function is a mean of each covariate to estimate models in a period t with vector $X_{1\times k}$ such that $\bar{X}_t = \frac{1}{N}\sum_{\mathrm{i}=1}^{N} X_{\mathrm{i},t}$, where $X_{\mathrm{i},t} \approx \bar{X}_t$.



(a) CUSUM Prob(e, $X\beta$) of Men (b) CUSUM Prob(e, $X\beta$) of Women

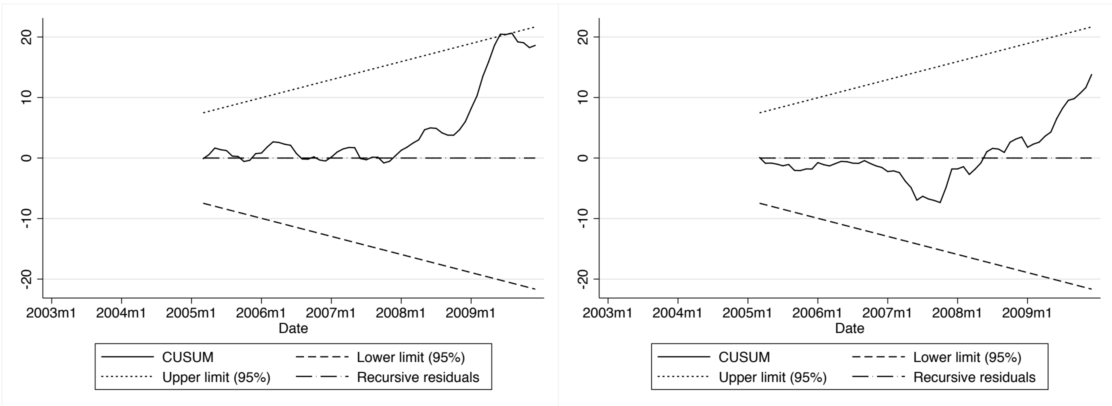
(c) CUSUMSQ Prob(e, $X\beta$) of Men (d) CUSUMSQ Prob(e, $X\beta$) of Women

Figure 3.2: Test for Parameter Stability Prob(e, $X\beta$)

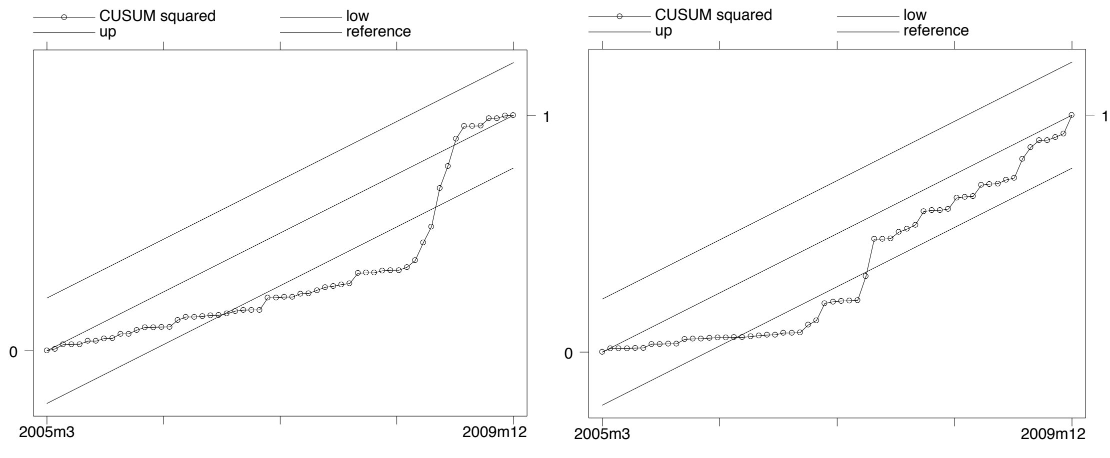
Figure 3.2, Figure 3.3, and Figure 3.4 present tests of stability to $s = \{\mathrm{e}, u, o\}$ in individual model $X\beta$. Figure 3.5, Figure 3.6, and Figure 3.7 present tests in model with aggregate effects Z.

Probability bounds for the CUSUM and CUSUMSQ have been tabulated according to grow linearly and are centered at 0. If the CUSUM or CUSUMSQ violate the bounds at any point, there is evidence of parameter instability. In this scenario, the errors will take a positive or negative sign and the sum tends to increase (in absolute value). On the other hand, if no structural change has occurred, the prediction errors should compensate each other and the cumulative sum should be close to zero.

To estimate CUSUM and CUSUMSQ we use an approximation, this is why an important issue is how boundaries are delimited. They are estimated as usual in these tests, but due to equations are approximation functions, their confidence bounds should be wider than tests do estimate. Even so, they are used as reference.



(a) CUSUM Prob$(u, X\beta)$ of Men (b) CUSUM Prob$(u, X\beta)$ of Women



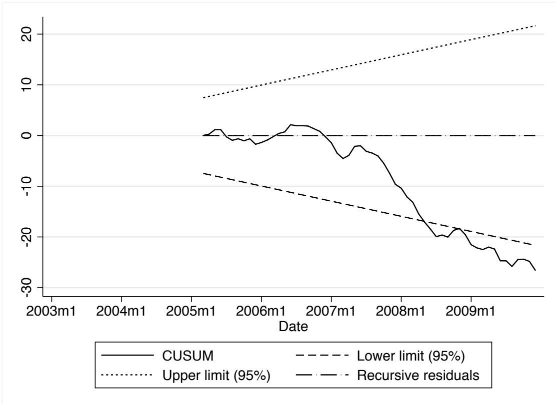
(c) CUSUMSQ Prob$(u, X\beta)$ of Men (d) CUSUMSQ Prob$(u, X\beta)$ of Women

Figure 3.3: Test for Parameter Stability Prob$(u, X\beta)$

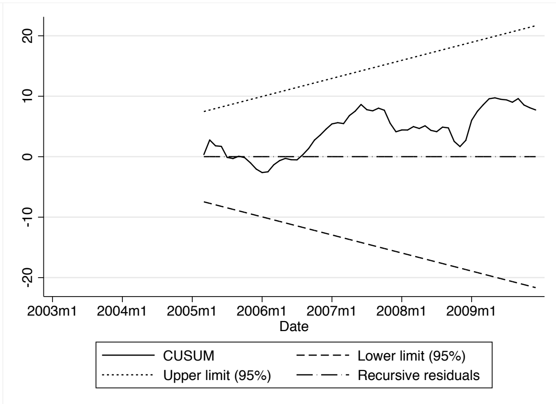
Using the CUSUM test, we can observe when the structural change occurs. However, the CUSUM test is not very powerful. Even though the structural change clearly takes place in a period, the null hypothesis is often accepted. Conversely, the CUSUMSQ test is too powerful but we cannot know the period when the structure changes (Perron, 2006).

Graphs of labor force status against individual and household characteristics suggest structural stability, although something may have changed around the time of the labor force status at the end of the analyzed period. Considering the CUSUM test of individual models, equations estimated for men are less stable for $s = o$, as Figure 3.4a presents, whereas equartions for women with $s =$ e as dependent variable show certain evidence of instability since early 2009, according to Figure 3.2b. CUSUMSQ illustrates some evidence of non-random deviation (but not necessarily this is an evidence of structural break) when $s = u$ both women and men. Particularly, there is a wider period of instability between 2007 and 2008 in men's labor force status, with shifted and persistent deviations.

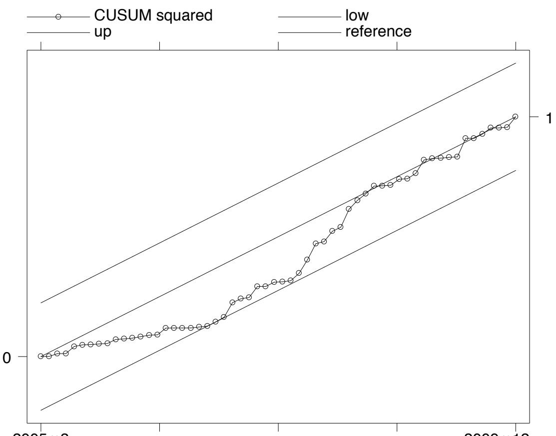
About evolution of stability, to estimate $s = u$, CUSUM evolution is more fluctuating both men and women since mid-2008. On the other hand, CUSUMQ presents a little non-random deviation when $s =$ e around 2008 both men and women.



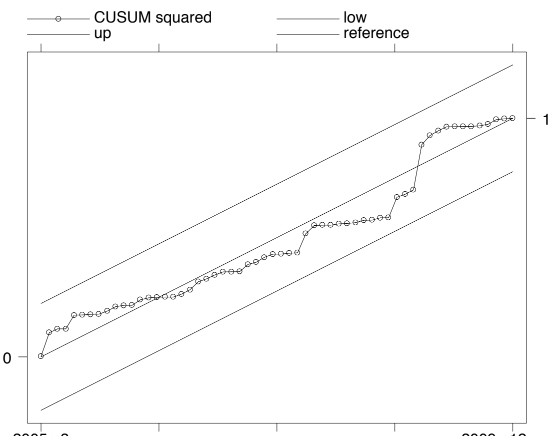
(a) CUSUM Prob($o, X\beta$) of Men



(b) CUSUM Prob($o, X\beta$) of Women



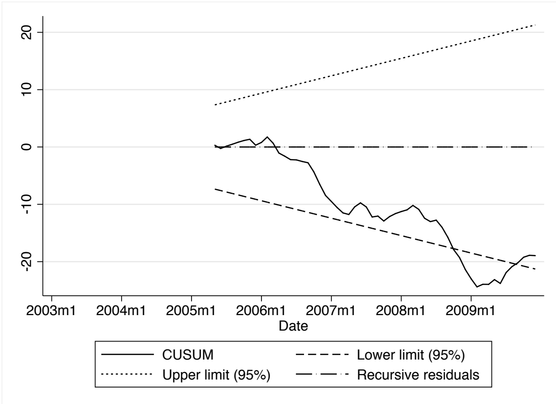
(c) CUSUMSQ Prob($o, X\beta$) of Men



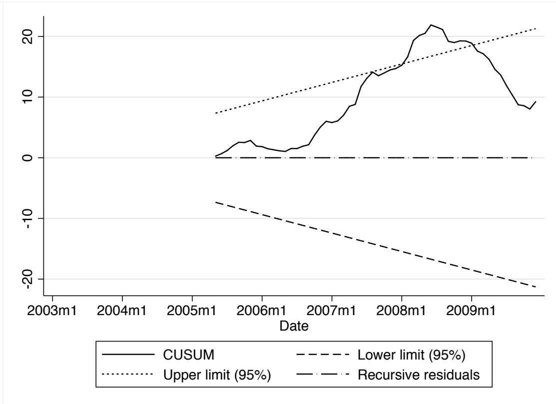
(d) CUSUMSQ Prob($o, X\beta$) of Women

Figure 3.4: Test for Parameter Stability Prob($o, X\beta$)

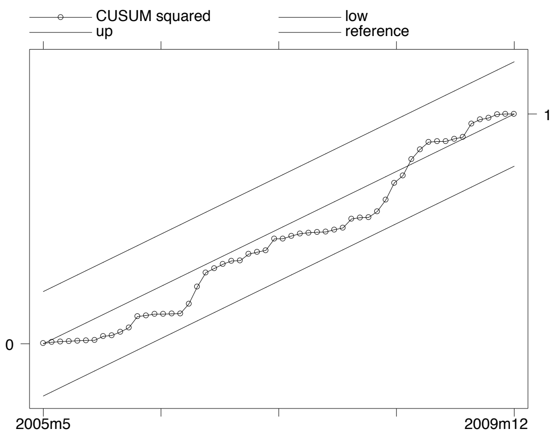
Note that the graphs start in May 2005, since recursive estimation begins using the first $k + 1$ data points.



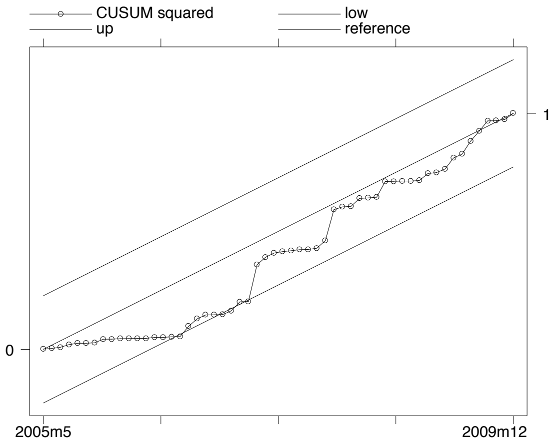
(a) CUSUM Prob(e, Z) of Men



(b) CUSUM Prob(e, Z) of Women



(c) CUSUMSQ Prob(e, Z) of Men



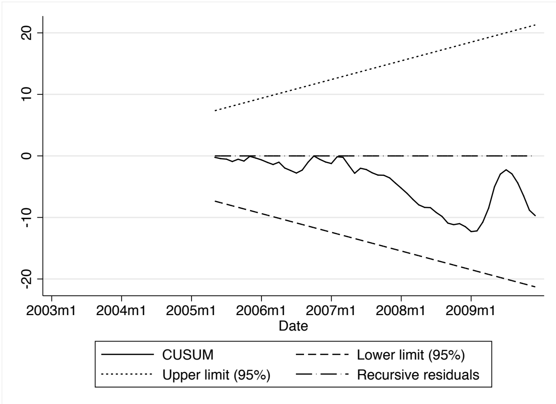
(d) CUSUMSQ Prob(e, Z) of Women

Figure 3.5: Test for Parameter Stability Prob(e, Z)

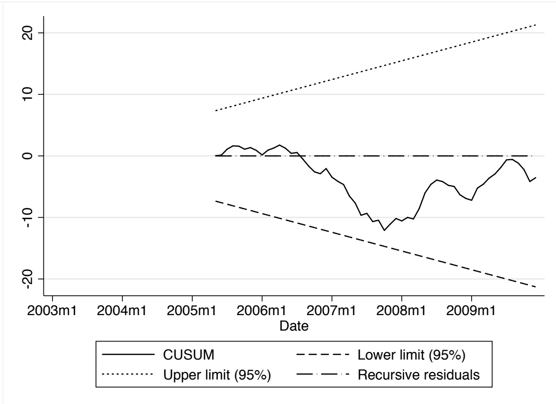
Graphs of labor force status against individual, household and market characteristics suggest, again, parameter stability. Thus, there are more signs of stability in predictive models with aggregate effects than individual models. Even so, according to CUSUM, $s = \mathrm{e}$ equations, both men and women, are more unstable, marked at last of period (since mid 2007 to women, and since early 2009 to men, according to Figure 3.5).

By CUSUMSQ, model estimated for $s = u$ to men presents more evidence of non-random deviations, with a long period of deviations between 2007 and the end of 2008, as Figure 3.6c illustrates. This fact proves more variance (CUSUMSQ has power for changing variance), similar situation for $s = o$ at last 2008. Models to women do not present evidence of non-random deviations in any labor force status according to CUSUMSQ, and adding CUSUM results, is a strong evidence of stability.

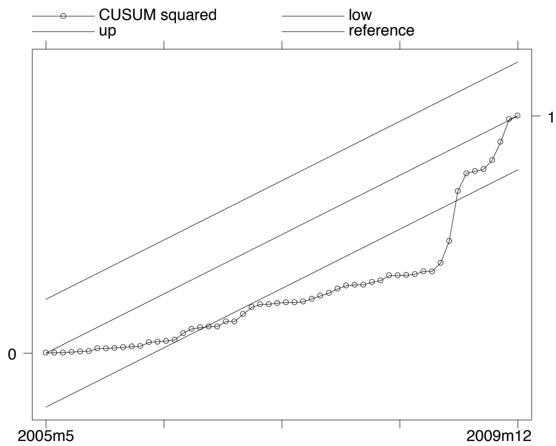
About evolution of stability, CUSUM test illustrates more fluctuations to all of three labor force status, especially since the mid of analyzed period. CUSUMSQ, by the other hand, shows much less variance in all equations.



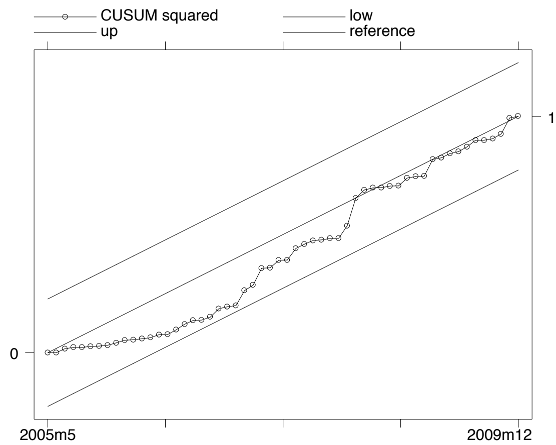
(a) CUSUM Prob(u, Z) of Men



(b) CUSUM Prob(u, Z) of Women



(c) CUSUMSQ Prob(u, Z) of Men

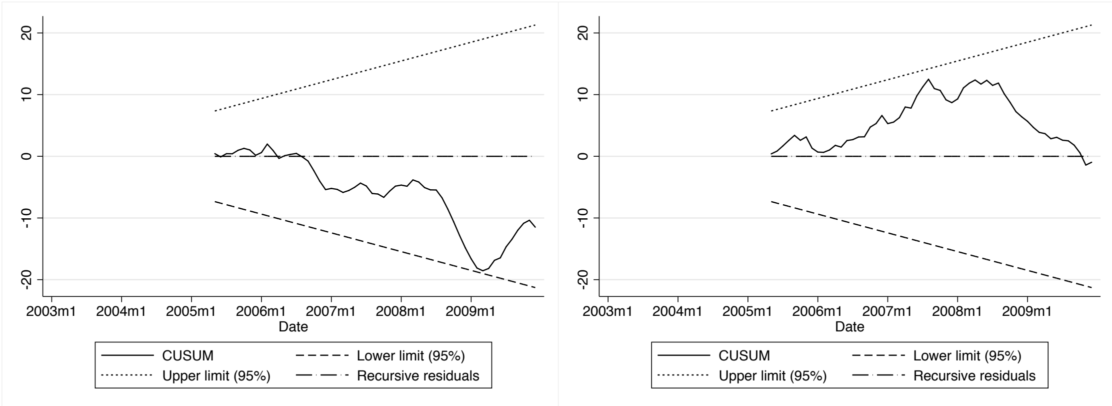


(d) CUSUMSQ Prob(u, Z) of Women

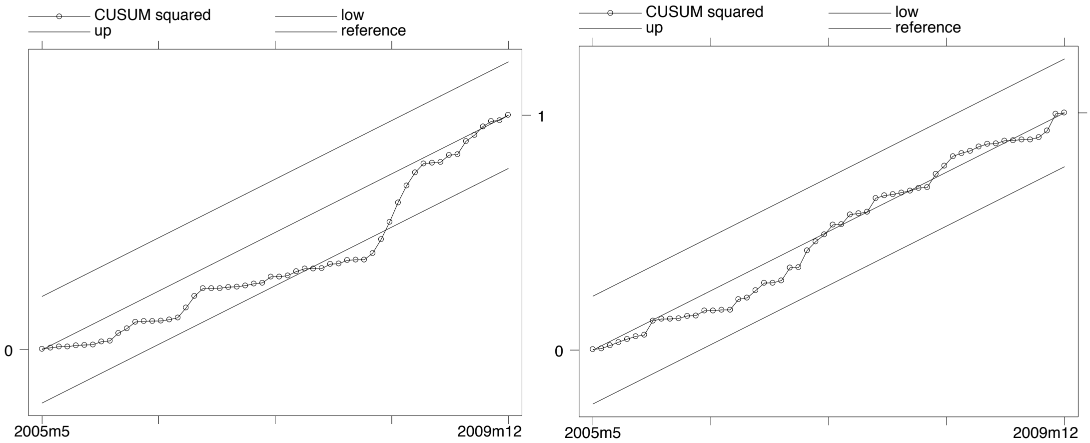
Figure 3.6: Test for Parameter Stability Prob(u, Z)

Neither case presents evidence of a structural change in estimated models of labor force status, since the cumulative sum stays within the confidence interval of a zero sum both CUSUM and CUSUMSQ. In all cases, we failed at finding evidence of a structural change (considering the fact that bounds are underestimate) and non-random deviations in amidst of a random process. Models with aggregate effects, however, present less signs of instability and random evolution, although both individual models and with aggregate effects present evidence of instability (though random deviations according to CUSUMSQ) at the end of analyzed period of time, coinciding with Subprime mortgage crisis.

The finding of instability may guide the selection of the window of data to be used for model estimation and forecasting (or lead to the use of rolling windows of observations to allow for gradual change). Because there is no evidence of structural change, it is possible to estimate counterfactuals directly for the time period selected by forecasting. So, "What would have happened to the labor force situation s, if it had not been exposed to a change in the methodology of measurement?".



(a) CUSUM Prob(o, Z) of Men (b) CUSUM Prob(o, Z) of Women



(c) CUSUMSQ Prob(o, Z) of Men (d) CUSUMSQ Prob(o, Z) of Women

Figure 3.7: Test for Parameter Stability Prob(o, Z)

To implement forecasting, $\widehat{\beta}$, $\widehat{\lambda}$, and $\widehat{\mu}$ are estimated in 2003-2009 monthly data (NES). Also, for the implementation of the model with aggregate effects, estimators are adjusted by δ. Then, they are used in the NNES 2010-2013 monthly data to estimate $\widehat{P}_t^s$, separately, individual model and a model with aggregate effects[7]. $\widehat{P}_t^s$ is later compared with observed proportion of being in any labor force status s.

[7]To Out-of-Sample forecast, it is not considered Reference 2009 NNES period due to variation in this period of result considering subsequent data, and the fact that is a period of reference with no official results. Reference 2009 NNES period is used only to compare the variation between NES and NNES in analysis of results.

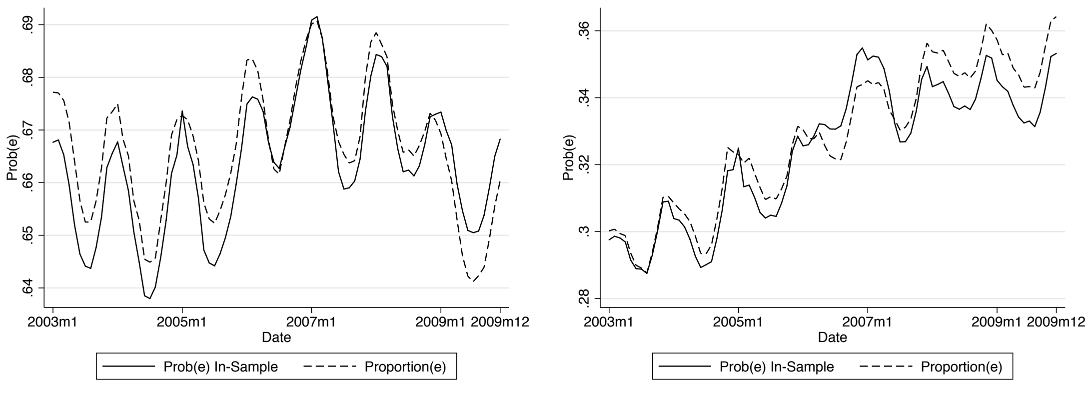
Chapter 4

Results

This chapter presents the main findings of labor force status evolution in the Chilean labor market from 2003 to 2013. For that, the first section presents the labor market fluctuations during the NES period, using an In-Sample analysis. The second section analyses Out-of-Sample forecasting results for the NNES period. As last, we discuss results contrasting the findings with prior evidence.

4.1 Labor Force Fluctuations 2003-2009

This section presents results of counterfactual probabilities by gender for the NES period. Figure 4.1, Figure 4.2, and Figure 4.3 present results for the individual model $X\beta$. Figure 4.4, Figure 4.5, and Figure 4.6 present results of the model with aggregate effects Z (adjusted).



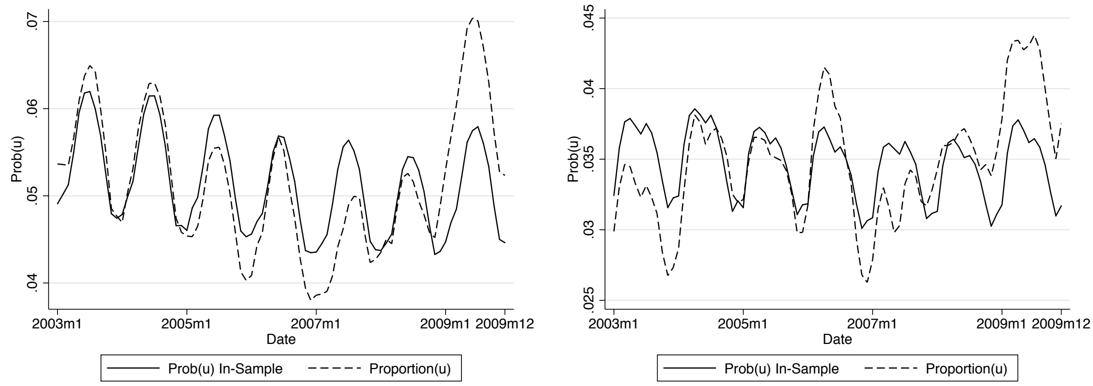
(a) $s = (\text{e}, X\beta)$ to men

(b) $s = (\text{e}, X\beta)$ to women

Figure 4.1: $s = (\text{e}, X\beta)$ 2003-2009, predicted and observed

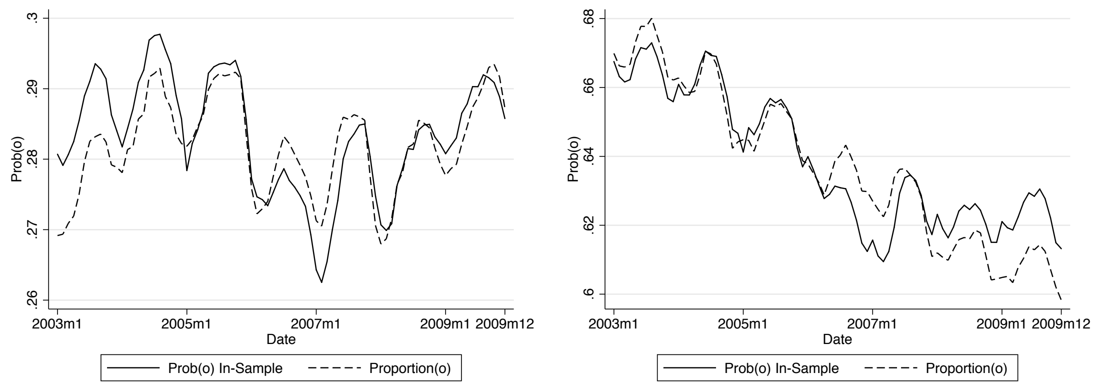
As all of these figures present, both models capture cyclicality of probabilities and perform well with respect to accurately forecasting (following error measures shown in subsection 3.2.1).

Depending on sex, and considering all the NES period, probabilities estimated to men are less accurate (also following the RMSE and Theil's U), but estimated probabilities to women are less adjusted at the end of period. This is true considering both estimated models, and it is more evident in the probability of being unemployed and to be out of labor market.



(a) $s = (u, X\beta)$ to men (b) $s = (u, X\beta)$ to women

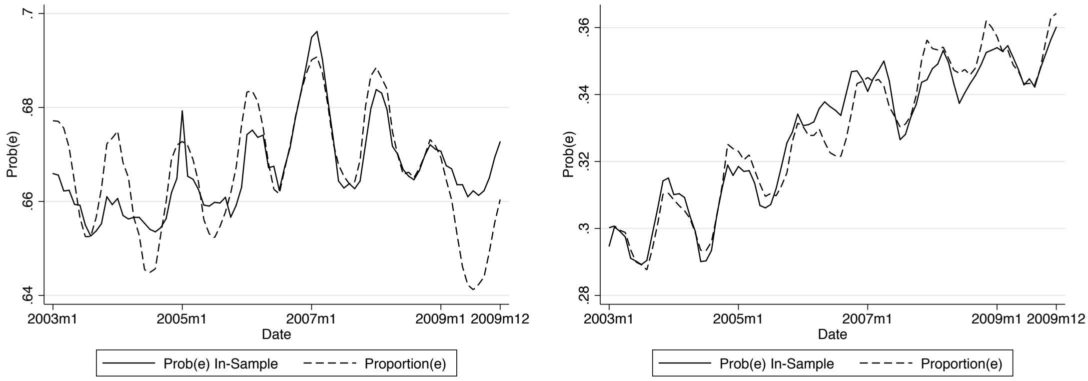
Figure 4.2: $s = (u, X\beta)$ 2003-2009, predicted and observed



(a) $s = (o, X\beta)$ to men (b) $s = (o, X\beta)$ to women

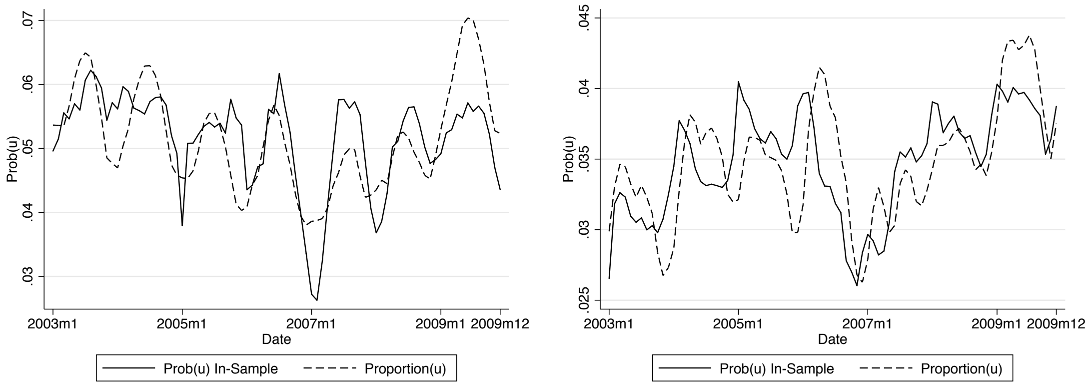
Figure 4.3: $s = (o, X\beta)$ 2003-2009, predicted and observed

The counterfactual probability of being unemployed shows less accuracy than in the other two labor force statuses, both for men and women, as observed in Figure 4.1 and Figure 4.5, following their higher values of Theil's U. On the one hand, both to men and women, counterfactuals estimated with individual models tend to underestimate these probabilities at the end of considered period. On the other, counterfactuals estimated using model with aggregate effects present some problems to predict unemployment in all In-Sample period.



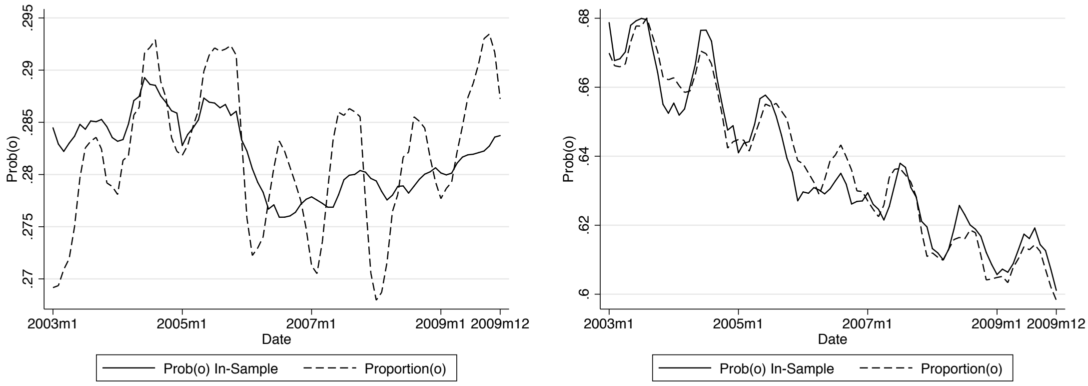
(a) $s = (\mathrm{e}, Z)$ to men (b) $s = (\mathrm{e}, Z)$ to women

Figure 4.4: $s = (\mathrm{e}, Z)$ 2003-2009, predicted and observed



(a) $s = (u, Z)$ to men (b) $s = (u, Z)$ to women

Figure 4.5: $s = (u, Z)$ 2003-2009, predicted and observed



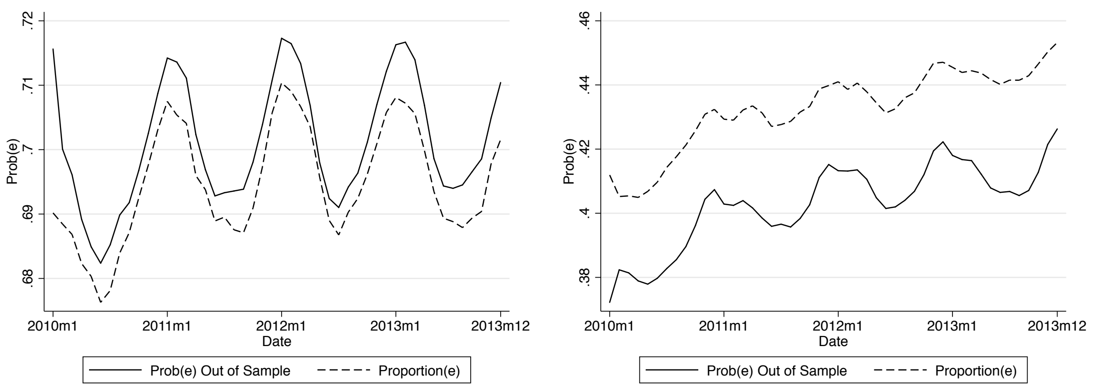
(a) $s = (o, Z)$ to men (b) $s = (o, Z)$ to women

Figure 4.6: $s = (o, Z)$ 2003-2009, predicted and observed

An important fact about the In-Sample counterfactual performance is that the probability of being out of labor market estimated using a model with aggregate effects to men presents problems throughout all period. So, at the beginning, it tends to overestimate probability, and since the mid of the period it tends to underestimate probability of being out of labor market, with less cyclicality than the observed proportion.

4.2 Fluctuations Since 2010

One defined the precision and adjusted (if needed) the counterfactual probabilities, and their In-Sample evolution, results of Out-of-Sample counterfactuals are analyzed. Figure 4.7, Figure 4.8, and Figure 4.9 present results for individual models $X\beta$. Then, Figure 4.10, Figure 4.11, and Figure 4.12 present results for aggregate effects models Z (predictions also adjusted using δ of In-Sample estimation).



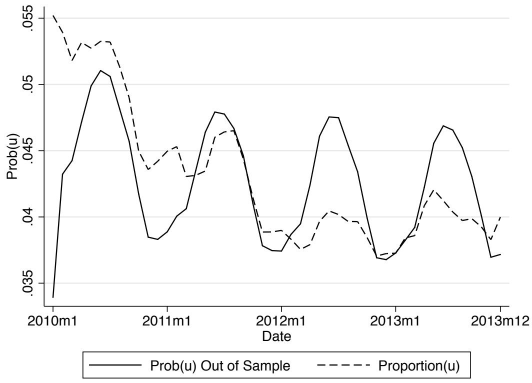
(a) $s = (\text{e}, X\beta)$ to men (b) $s = (\text{e}, X\beta)$ to women

Figure 4.7: $s = (\text{e}, X\beta)$ 2010-2013, predicted and observed

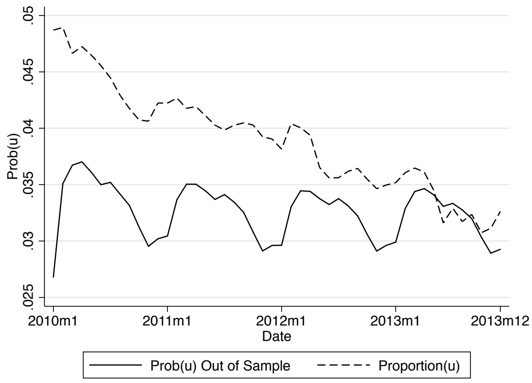
As all these figures present, both models capture cyclicality of observed proportions. In fact, counterfactual predictions are more cyclical than observed proportions. By model, the individual model captures more cyclicality than the model with aggregate effects, particularly to predict being out of labor market.

When assessing evolutions of counterfactuals by sex, predictions estimated to men are more cyclical, when comparing both models. Cyclicality is clearer in the prediction of employment (Figure 4.7a and Figure 4.10a) and unemployment (Figure 4.8a and Figure 4.11a).

The-counterfactual with more variation respect to the observed proportion is the probability of being unemployed. More marked trends are observed in the prediction for women, both for the individual model and for the model with aggregate effects, as presented by Figure 4.8b and Figure 4.11b. Also, a difference between the observed proportion and the estimated probability is more evident in predictions made by the model with aggregate effects, as presented by Figure 4.11. On the other hand, the probability of being out of



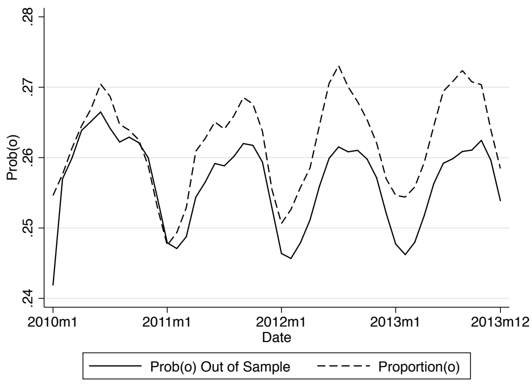
(a) $s = (u, X\beta)$ to men



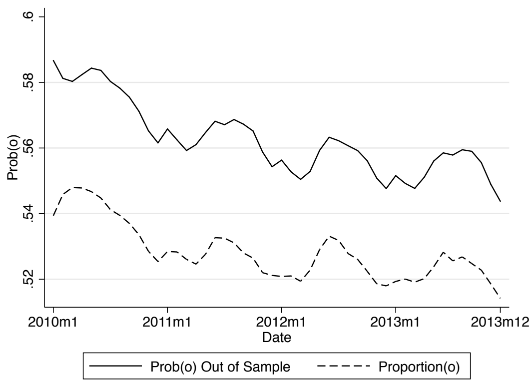
(b) $s = (u, X\beta)$ to women

Figure 4.8: $s = (u, X\beta)$ 2010-2013, predicted and observed

the labor market is more stable throughout time, particularly the probability estimated with the model with aggregate effects, presented in Figure 4.12.

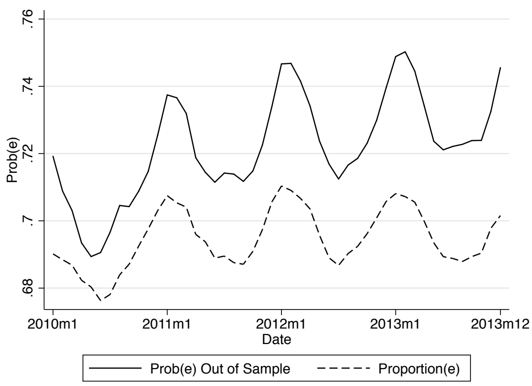


(a) $s = (o, X\beta)$ to men

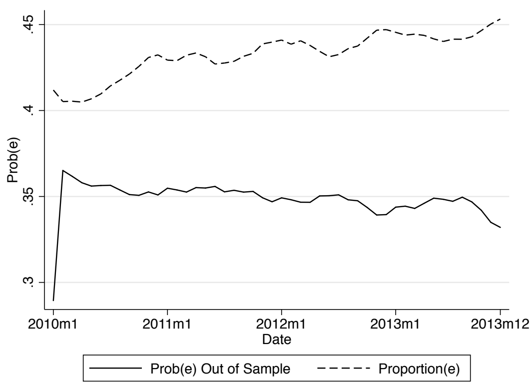


(b) $s = (o, X\beta)$ to women

Figure 4.9: $s = (o, X\beta)$ 2010-2013, predicted and observed

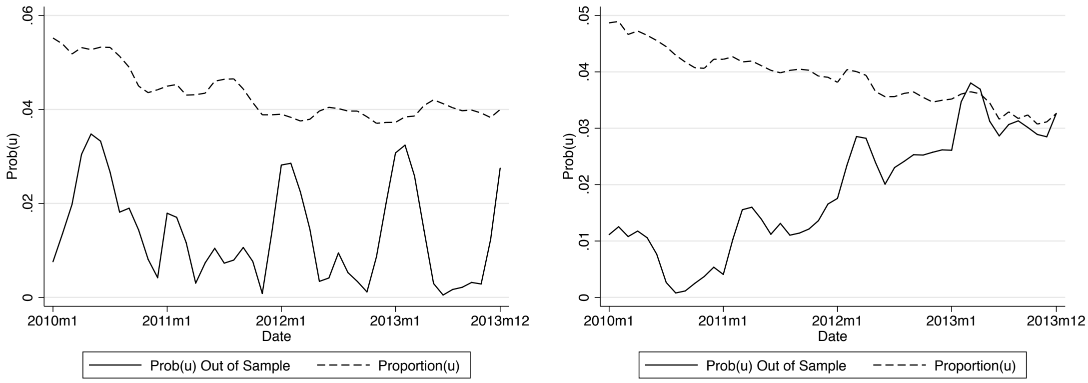


(a) $s = (\mathrm{e}, Z)$ to men



(b) $s = (\mathrm{e}, Z)$ to women

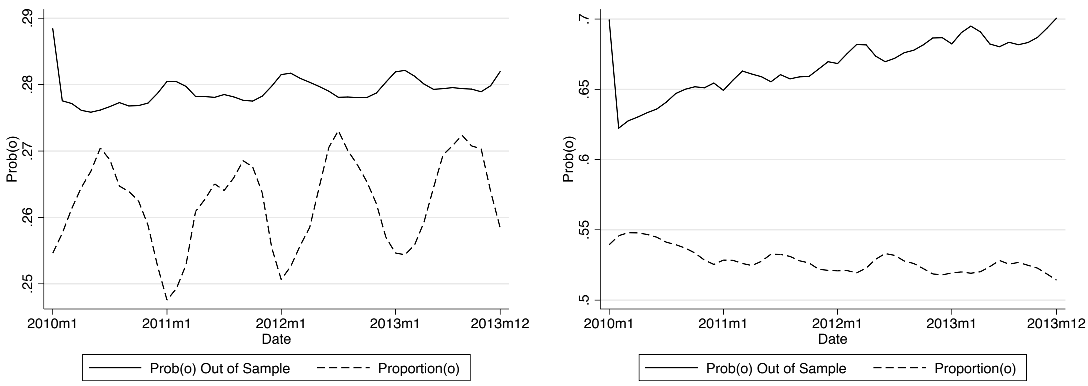
Figure 4.10: $s = (\mathrm{e}, Z)$ 2010-2013, predicted and observed



(a) $s = (u, Z)$ to men

(b) $s = (u, Z)$ to women

Figure 4.11: $s = (u, Z)$ 2010-2013, predicted and observed



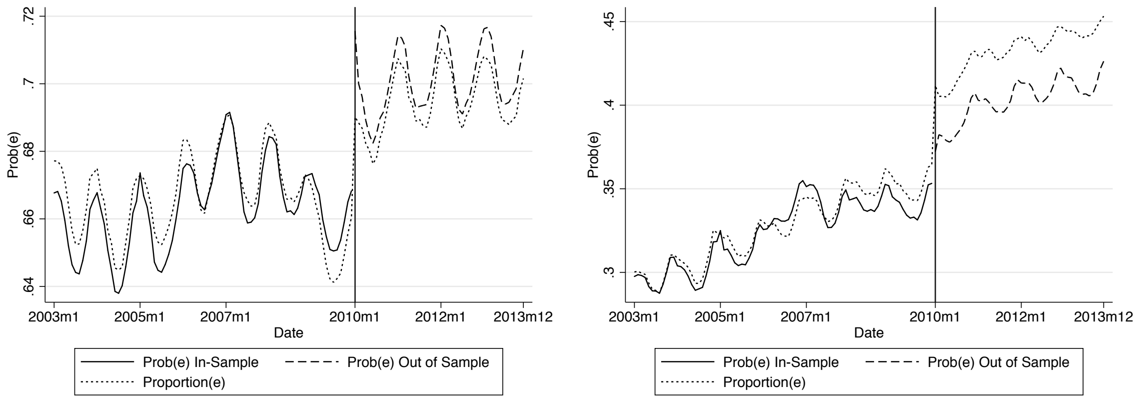
(a) $s = (o, Z)$ to men

(b) $s = (o, Z)$ to women

Figure 4.12: $s = (o, Z)$ 2010-2013, predicted and observed

4.3 Labor Market Dynamics 2003-2013

Previous sections analyzed performance of In-Sample and Out-of-Sample counterfactuals. Now, this section undertakes counterfactuals exercises by assuming what would have happened to labor market conditions analysis during 2003-2013 assuming that the NES had remained in use. Figure 4.13, Figure 4.14, and Figure 4.15 present performance of individual model $X\beta$ with respect to observed proportion of being in $s = \{\text{e}, u, o\}$. Also, Figure 4.16, Figure 4.17, and Figure 4.18 present performance of model with aggregate effects Z with respect to observed proportion of being in s.

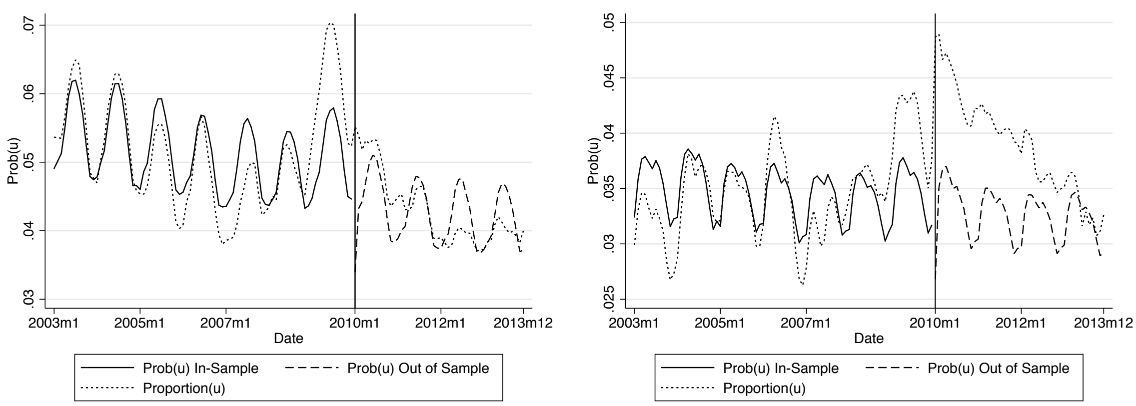


(a) $s = (\mathrm{e}, X\beta)$ to men (b) $s = (\mathrm{e}, X\beta)$ to women

Figure 4.13: $s = (\mathrm{e}, X\beta)$ 2003-2013, predicted and observed

Overall the figures illustrate that counterfactuals capture cyclicality of observed proportions measured in NES, which are remained in the NNES period. Also, they illustrate clearer trends of labor force status —and more in agreement with evolution of observed trends prior to 2010— than results obtained by the NNES. It is noteworthy that counterfactuals present smaller probabilities to be unemployed, but these probabilities follow prior trend (in particular to consider differences between the NES and the NNES explained in section 2.3).

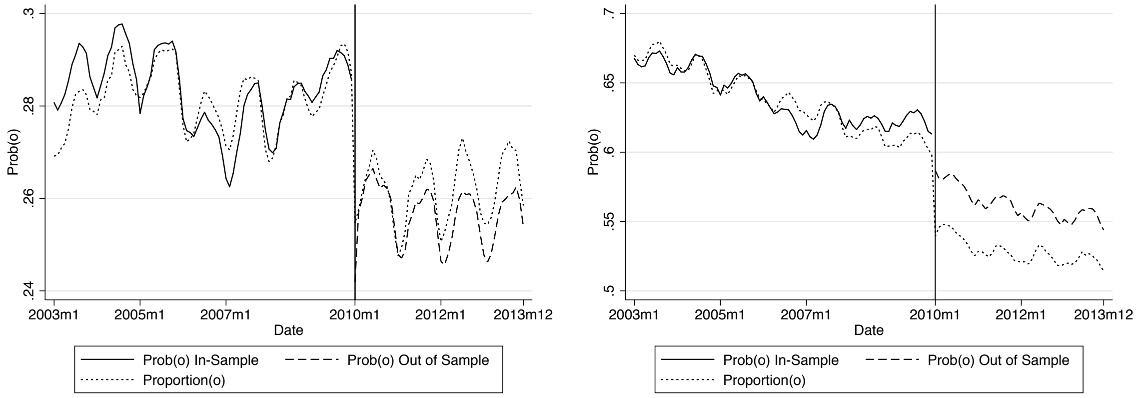
According to counterfactuals estimated by individual model, constant increasing in labor force participation is explained by higher male employment than female's, and by lower unemployment in men. Similar to counterfactuals estimated using model with aggregate effects, where increases in labor force participation are explained by higher male employment than female's, and by smaller unemployment in men.



(a) $s = (u, X\beta)$ to men (b) $s = (u, X\beta)$ to women

Figure 4.14: $s = (u, X\beta)$ 2003-2013, predicted and observed

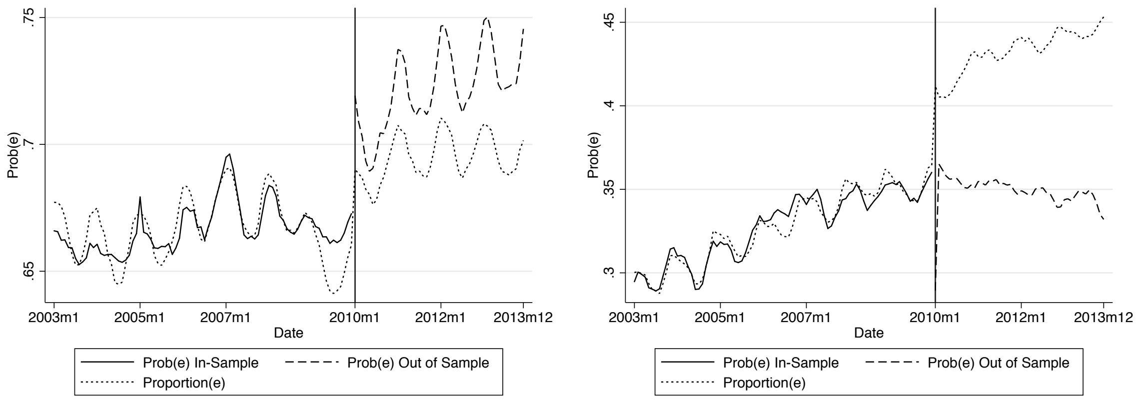
Following the predictions made using individual models, predicted probabilities of being in s illustrate an uninterrupted increase in labor force participation since last 2009, although is less pronounced than observed proportions show (considering the NNES results). Differences are particularly true in women, for whom probability of being employed would have been lower if survey (and method) had not changed; and probability of being out of labor market is larger than new measurement presents, to men. Also, probability of being unemployed keeps fluctuating trends observed in the NES period, both men and women.



(a) $s = (o, X\beta)$ to men (b) $s = (o, X\beta)$ to women

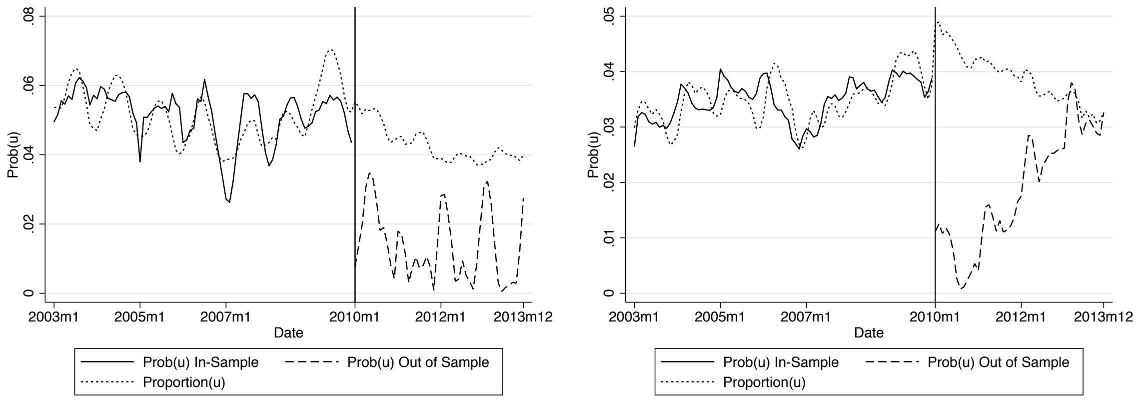
Figure 4.15: $s = (o, X\beta)$ 2003-2013, predicted and observed

Counterfactuals estimated using model with aggregate effects, on the other hand, illustrate more difference respect to observed proportions as same as less cyclicality. First, increases in labor force participation are smaller than observed in counterfactuals estimated by individual model and observed proportions, particularly since 2010. Even though probabilities of being employed and out of labor market in the NNES period follow prior trends —both to men and women—, probability of being unemployed sharp falls at the beginning of the NNES period. Also, this probability falls more than observed proportion, and then, it increases constantly, but more cyclically.



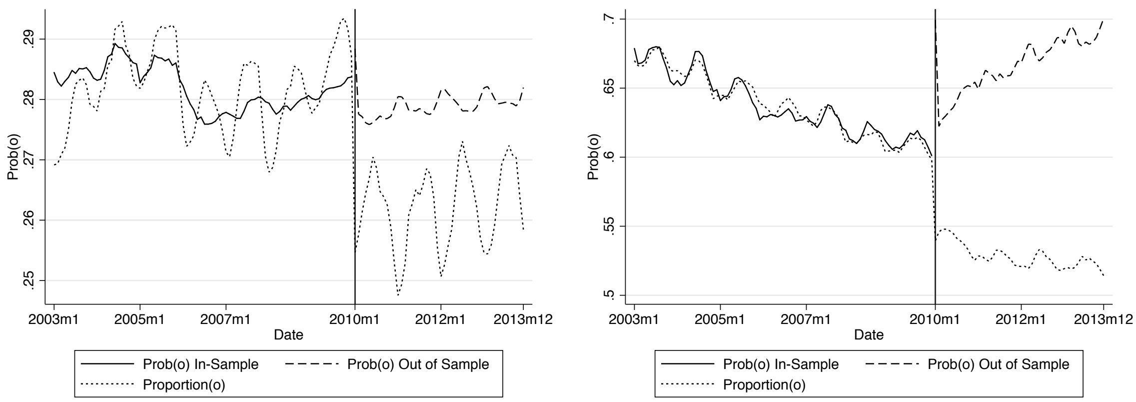
(a) $s = (\mathrm{e}, Z)$ to men (b) $s = (\mathrm{e}, Z)$ to women

Figure 4.16: $s = (\mathrm{e}, Z)$ 2003-2013, predicted and observed



(a) $s = (u, Z)$ to men (b) $s = (u, Z)$ to women

Figure 4.17: $s = (u, Z)$ 2003-2013, predicted and observed



(a) $s = (o, Z)$ to men (b) $s = (o, Z)$ to women

Figure 4.18: $s = (o, Z)$ 2003-2013, predicted and observed

4.4 A Brief Discussion

After analyzing results obtained through counterfactuals and their performance in relation to observed proportions of s, it becomes necessary to ask whether the research question can be answered. Also, results should deal with the hypotheses that support this work.

First, using only counterfactual findings it would not have made possible to deal with the evolution of the labor market in the mid-term. But in short-run, there is some evidence that sheds light on a positive dependence on business cycles, particularly considering market recovery after the Subprime Crisis. This evidences highlights that there has been no sizable structural change in economy (Bravo et al., 2005; Cowan et al., 2005; Landerretche, 2007).

Second, tools used to measure labor market performance have a significant impact on the explanation of labor force status evolution, and our results strengthen that conclusion. Two main ideas arise from this. Results illustrate that market dynamics are indeed more cyclical using the NES than the NNES method. Thus, post-2010 fluctuations are attributable to changes in the measurement of the labor force status, following Marcel and Naudon (2016) results. Results, also, confirm differences between the NES and the NNES from the Reference Series 2009 presented in section 2.3. Differences highlight that the NES underestimates male employment and overestimates female employment in respect to the NNES. In particular, and because women are worse affected by job insecurity and working less hours than they would like, differences between measures are an important issue. Another key difference is that the NES underestimates unemployment when is compared to the NNES, because requirements to be classified as unemployed are more demanding in the NES, as explained Table 2.1. Even so, the lack of cyclicality of the NNES results affects, in particular, measures of unemployment. So, unemployment evolution post-the Subprime Crisis has remained fairly stable; and its recovery has been weaker than observed data illustrates. Moreover, as the probability to be classified as employed differs by sex, being considered out of labor market is underestimated both for men and women, being more marked for women, and follows the overestimated probability of being employed.

There are two important points to remark. The unemployed people is a tiny proportion of the working-age people, so it is reasonable to have empirical problems. Also, the decreasing trend and the cyclicality of counterfactuals of being out of labor market (especially among women) is consistent with the literature, although this work considers it as a stock. Additionally, the performance of models estimated for men is worse than models estimated for women, according to error measures. A tentative explanation for this corresponds to the models themselves, *i.e.* it may exist omitted-variable bias because male labor market insertion is not the same as female's.

Third, aggregate effects are not the best instrument to add labor market conditions into an analysis of labor force status as expected. Probably, it is necessary to think about other options to measure these effects. So, how this thesis measures aggregate effects (as conditions at the same period that one person is in front of labor market) may be insufficient to explain s.

To talk about the strength of our results, it is important to consider that the exercise to estimate counterfactuals is only possible due to Assumptions 1, 2, and 3, and neither supply shocks nor demand for labor have been played major roles in the Chilean economy[1]. So, counterfactual results are robust and important as a first approximation.

Dealing with a large share of the research question there are some data weaknesses to consider. The most important one is the trouble to make an analysis of fluxes considering the methodology change. This would have been an interesting complement for our results. In addition, it is not possible —at the moment— to conduct analysis about Separation

[1]Both models fulfill the two key assumptions of forecasting, so is possible to assume forecasting captures regularities of chilean labor market, and later results present good fitting In-Sample.

or Job-finding rate. Also, stock analysis does not consider tensions between the labor supply and demand underlying to fluctuations found in this work. Likewise, analysis of unemployment would be more complete if consider the opportunity cost of employment.

Lastly, this work leaves some unresolved matters that may be considered in further research, such as: Doing forecasting using a wider period of the NES data to make an analysis more complete; making the inverse exercise through forecasting the NNES parameters into the NES data (or "What happened with labor force status evolution if the NNES would be collected since earlier?"); and, testing other measurement mechanisms for aggregate effects, maybe considering lags of variables estimated as effects.

Concluding Remarks

The main aim of this thesis was to examine the labor market dynamics in Chile between 2003 and 2013. That purpose involved dealing with data comparability problems, for which methods based on models to estimate counterfactuals of probabilities of being in a labor force status (as employed, unemployed, or out of labor market) were used. To estimate counterfactuals, two models were proposed: the first model considers individual characteristics (both personal and household conditions), whereas the second model considers individual characteristics and aggregate effects, defined as the probability of an individual in working-age population of being in a labor force status due to pressure of the group $g(\mathrm{i})$ in a period of time. Then, probabilities were estimated for the period 2003-2009 (years where the NES was collected), then, estimation parameters were used to forecast predictions for the period 2010-2013 (where the NNES has been collected). From this work, four main results arise from the analysis.

At first, and answering the research question, labor market dynamics would have been steady between 2003 and 2013, despite Subprime Crisis if measurement method of labor force status had not changed. The crisis implied a decrease in employment and an increase in unemployment although they were no sharply. *Ex-post* recovery, in the same way, is more gradual and stable regarding labor market conditions prior to the crisis, as observed NNES results indicate. Comparing the observed proportion of people in s and the predicted probability of being in the same status s, trends are different depending on gender. Thus, to women, the NNES overestimates probability of being employed and underestimates probability of being out of labor market due to the one-hour criterion in the definition of employment emphasize irregular work, which affects importantly to women. To men, the NNES slightly underestimates the probability of being employed and slightly overestimates the probability of being out of the labor market. The NNES overestimates, both to men and women, the probability of being unemployed because of the less exhaustive requirements to be considered in it. An important issue is the lack of cyclicality of the NNES results respect to the NNES, particularly when classified as unemployed.

Secondly, differences between observed data and counterfactuals present far-reaching consequences to analyze the NNES results. Fluctuations in labor market dynamics since 2010 are attributable to changes in the measurement of labor force status. An explanation of the evolution of the labor force status should consider participation or, at least, the relation between employment and to be out of labor market. This is because participation rate is much more precise thermometer of labor market performance due to lack of cyclicality in the measure of unemployment in the NNES. This could illustrate that if the key consequence of Asian Crisis was the persistent higher unemployment rates, the most important consequence of Subprime Crisis is the persistent movement between employment and inactivity.

Thirdly, and from an overall perspective, it is clear that fluctuations in labor market conditions depend on business cycles and tools used to measure them. On the one hand, according to results, the Chilean labor market evolution depends on business cycles in the short-run, which is consistent to labor market dynamics during the late nineties and 2000s (Bravo et al., 2005; Cowan et al., 2005; Landerretche, 2007; Naudon & García, 2012; Sehnbruch, 2006), and early evidence about last years (Marcel & Naudon, 2016). So, the grow in probability of being employed and the decline in probability of being unemployed illustrate that nowadays, the Chilean labor market is less damaged than during the Asian Crisis, when it experienced a persistent unemployment rate. On the other hand, both results of Reference Series 2009 (comparing the NES and the NNES) and counterfactuals illustrate that NES results are indeed more cyclical than NNES results. It seems that using the one-hour criterion in the definition of employment, trying to cover all types of employment implies a lack of cyclicality of evolution of labor force status.

Finally, counterfactuals estimated using an individual model and a model including aggregate effects capture dynamics of labor market in the NES period, thus is expected their suitable running in the NNES period (Out-of-Sample). Also, assessing goodness-on-fit models In-Sample, the individual model performs better than the model with aggregate effects. So, to consider aggregate effects in further works its measurement must be improved. Also, to assess counterfactuals performance in a wider context, they must be estimated using a different and a wider period of time. Or the inverse process: this is another reason to estimate counterfactuals of the NNES performance if this survey would have been collected since earlier (e.g. early 2000s).

Bibliography

Acemoglu, D. (2008). *Introduction to Modern Economic Growth.* Princeton: Princeton University Press.

Ammermueller, A., & Pischke, J.-S. (2006). *Peer Effects in European Primary Schools: Evidence from PIRLS.* Bonn: Discussion Paper No. 2077. Institute for the Study of Labor (IZA).

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

Angrist, J., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics.* Princeton: Princeton University Press.

Bandiera, O., Barankay, I., & Rasul, I. (2010). Social incentives in the workplace. *Review of Economic Studies*(77), 417–458.

Barro, R., & Sala-i Martin, X. (2004). *Economic Growth.* Cambridge, MA: MIT Press.

Blanchard, O., & Katz, L. (1992). Regional Evolutions. *Brookings Papers on Economic Activity*, *23*(1), 1–76.

Blau, F., & Kahn, L. (2013). The feasibility and Importance of Adding Measures of Actual Experience to Cross-sectional Data Collection. *Journal of Labor Economics*, *31*(S1), 17–58.

Boeri, T., & Van Ours, J. (2008). *The Economics of Imperfect Labor Markets.* Princeton: Princeton University Press.

Bravo, D., Ferrada, C., & Landerretche, O. (2005). *The Labor Market and Economic Cycles in Chile.* Universidad de Chile, Santiago.

Calvó-Armengol, A., & Jackson, M. (2004). The Effects of Social Networks on Employment and Inequality. *American Economic Review*, *94*(3), 425–454.

Cameron, C., & Trivedi, P. (2010). *Microeconometris using Stata.* College Station: Stata Press.

Cowan, K., Micco, A., Mizala, A., Pagés, C., & Romaguera, P. (2005). *Un Diagnóstico del Desempleo en Chile.* Santiago: Centro de Microdatos, Universidad de Chile.

Cruces, G., & Galiani, S. (2005). Fertility and Female Labor Supply in Latin America: New Causal Evidence. *Labour Economics*, *17*, 565–573.

DANE. (2009). *Metodología Gran Encuesta Integrada de Hogares.* Bogotá: Departamento Administrativo Nacional de Estadística.

DANE. (2012). *Colombia: Gran Encuesta Integrada de Hogares. Metodología y cuestionario.* Bogotá: Comisión Económica para América Latina y el Caribe.

Ehrenberg, R., & Smith, R. (2009). *Modern Labor Economics. Theory and Public Policy.* Boston: Pearson Education.

Elliott, G., & Timmermann, A. (2008). Economic Forecasting. *Journal of Economic Literature*, *46*(1), 3–56.

Fleck, S., & Sorrentino, C. (1994). Employment and Unemployment in Mexico's Labor Force. *Monthly Labor Review*, *117*, 3–31.

Greene, W. H. (2011). *Econometric Analysis.* Upper Saddle River: Pearson Education.

Gutiérrez, R. (2009). La Recesión y la Medición de la Tasa de Desempleo de México Ante una Eventual Reforma Laboral. *Economía: teoría y práctica*(31), 43-72.

Hall, R. (2005). Job Loss, Job Finding, and Unemployment in the US Economy Over the Past Fifty Years. *NBER Macroeconomics Annual*, *20*, 101–137.

Hansen, B. (1992). Testing for Parameter Instability in Linear Models. *Journal of Policy Modeling*, *14*(4), 517–533.

Hansen, P. R., & Timmermann, A. (2012). *Choice of Sample Split in Out-of-Sample Forecast Evaluation.* Stanford University, Palo Alto.

Heckman, J., & Macurdy, T. (1980). A Life Cycle Model of Female Labour Supply. *The Review of Economic Studies*, *47*(1), 47–74.

Hussmanns, R., Mehran, F., & Varmā, V. (1990). *Surveys of Economically Active Population, Employment, Unemployment, and Underemployment: an ILO Manual on Concepts and Methods.* Geneva: International Labour Organization.

IBGE. (2015a). *Nota Técnica Principais diferenças Metodológicas Entre as Pesquisas PME, PNAD e PNAD Contínua.* Rio de Janeiro: Instituto Brasileiro de Geografía e Estatística.

IBGE. (2015b). *Pesquisa Nacional por Amostra de Domicílios. Síntese de Indicadores.* Rio de Janeiro: Instituto Brasileiro de Geografía e Estatística.

ILO. (2004). *Comparable Annual Employment and Unemployment Estimates.* Geneva: International Labour Organization.

INE. (2010). *Nueva Encuesta Nacional de Empleo. Manual Metodológico y Conceptual.* Santiago: Instituto Nacional de Estadísticas.

Inoue, A., & Kilian, L. (2005). In-Sample or Out-of-Sample Tests of Predictability: Which One Should we Use? *Econometric Reviews*, *23*(4), 371–402.

Kale, Y., & Doguwa, S. I. (2015). On the Compilation of Labour Force Statistics for Nigeria. *Central Bank of Nigeria Journal of Applied Statistics*, *6*(1), 183–198.

Landerretche, O. (2007). *Job Flows in Chile.* Santiago: Serie Documentos de Trabajo No. 240, Departamento de Economía, Universidad de Chile.

Marcel, M., & Naudon, A. (2016). *Transiciones Laborales y la Tasa de Desempleo en Chile.* Santiago: Banco Central de Chile. Documento de Trabajo No. 787.

Martin, G. (2000). Employment and Unemployment in Mexico in the 1990s. *Monthly Labor Review*, *123*, 3–18.

Mas, A., & Moretti, E. (2009). Peers at Work. *American Economic Review*, *99*(1), 112–45.

Márquez, C. (2015). Determinantes del Desempleo en las Urbes Mexicanas. Continuidades y Rupturas en el Periodo de Crisis. *Papeles de Población*, *21*(83), 101–134.

Naudon, A., & García, M. (2012). *Dinámica Laboral en Chile.* Santiago: Banco Central de Chile. Documento de Trabajo No. 659.

NBS. (2014). *Report of National Stakeholders' Workshop on The Review of Definition and Methodology for Computing Unemployment Statistics in Nigeria.* Abuja: National Bureau of Statistics.

NBS. (2015). *Unemployment/Under-employment Watch Q2 2015.* Abuja: National Bureau of Statistics.

OECD. (2017). *OECD Employment Outlook 2017.* Paris: Organisation for Economic

Co-operation and Development.

Perron, P. (2006). Dealing with Structural Breaks. In T. Mills & K. Patterson (Eds.), *Palgrave handbook of econometrics* (Vol. 1, pp. 278–352). Wiltshire: Palgrave Macmillan UK.

Pindyck, R., & Rubinfeld, D. (1998). *Econometric models and economic forecasts.* New York: McGraw-Hill.

Salazar, C. (2006). *Fortalezas y limitantes de la Encuesta Continua de Hogares ECH frente a estudios del Mercado laboral colombiano.* Bogotá: Departamento Nacional de Planeación.

Sehnbruch, K. (2006). *The Chilean Labor Market.* New York: Palgrave Macmillan.

Statistics NZ. (2016a). *Household Labour Force Survey: Revisions to Labour Market Mstimates.* Wellington: Statistics New Zealand.

Statistics NZ. (2016b). *Household Labour Force Survey. Summary of 2016 Redevelopment.* Wellington: Statistics New Zealand.

Topa, G. (2011). Labor Markets and Referrals. In B. A. Benhabib Jessie & M. Jackson (Eds.), *Handbook of Social Economics* (Vol. 1B, pp. 1193–1222). Elsevier.

Train, K. E. (2009). *Discrete Choice Methods with Simulation.* New York: Cambridge University Press.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Appendix A

Variable Descriptions

Appendix A contains the detailed definitions of variables used in analysis. Components of the vector of individual characteristics X are described in Table A.1. Also, in Table A.2 are described categories of discrete covariates.

An important issue is the fact that variable region of living was recoded to assure identification[1]. Clustering criteria is based in regional GDP per economic activity by region, data published by the Central Bank of Chile for the year 2014. Clustering is detailed in Table A.3.

Variable	**Definition**
Marital Status	Refers to a relation of domestic partnership and responsabilities
Schooling	Consider as formal educational stages, that act as signaling in labor market of qualifications and skills
Head of Household	Refers to a person who has the responsibility for governing a group that lives together, such as a family.
Age	Age of people, transformed into a discrete variable
Household Dependency ratio	Number of people who have to be supported financially by each person employed in a given population. Ratio between number of employed persons at household (at working-age) and the total number of people at household. It is used to measure the pressure on productive people at household.
Month of Survey	Month where was applied survey, used to controlling seasonality

Table A.1: Description of components of vector X

[1] There is an identification problem in more distant regions, since population is concentrated in the center of the country.

Variable	**Definition**
Marital Status	1 Married people
	0 Single, widow or divorced people.
Schooling	1 No education or less than Primary Education
	2 Completed Primary Education
	3 Some Secondary Education
	4 Secondary Education Graduated
	5 Some Higher Education
	6 Higher Education Graduated
Head of Household	0 Not head of Household
	1 Head of household
Age	1 Between 15 and 24 years old
	2 Between 25 and 29 years old
	3 Between 30 and 35 years old
	4 Between 35 and 45 years old
	5 Between 45 and 55 years old
	6 Between 55 and 65 years old

Table A.2: Description of Values of Variables to Compone Vector X

Recoding Region	**Official Region**
Northern Region. Their economical activity is principally Mining, Construction	Arica y Parinacota
	Tarapacá
	Antofagasta
	Atacama
Central-Coast Region. Their economical activity is based in Mining and Tertiary sector (as Personal Services, Telecommunication and Entertainment)	Coquimbo
	Valparaíso
Central-Agro Region. Economical activity are Agriculture, Manufacturing Industry and Personal Services	O'Higgins
	Maule
Manufacturing South Region. Activities are Manufacturing Industry, and Tertiary sector (Personal Services, Government and Entertainment)	Bío Bío
	La Araucanía
Southern Region. Economical activity are Manufacturing Industry, Construction and Government	Los Ríos
	Los Lagos
	Aysén
	Magallanes y de la Antártica Chilena
Metropolitan County	Región Metropolitana

Table A.3: Recoding Variable Region of Residence

Appendix B

Additional Regressions

Appendix B provides the results of probit models to dependent variable $\{s = \mathrm{e}, u, o\}$, both individual and with aggregate effects. Remember, probit models are estimated separately by sex. Years included are 2003-2009, which form the pooling with NES data to estimate parameters to forecast. Also, each regression controls for seasonality (monthly). All predictors are estimated at their mean value.

Order of regressions is as follows. First, Table B.1, Table B.2, and Table B.3 present estimates using individual models to women and Table B.4, Table B.5, and Table B.6 to men. Then, are presented regressions using aggregate effects and interaction terms. To $s = \mathrm{e}$, Table B.7, Table B.8, and Table B.9 are models to women, and Table B.16, Table B.17, and Table B.18 to men. Same order to regressions that consider $s = u$ with Table B.10, Table B.11, and Table B.12 to women, and Table B.19, Table B.20, and Table B.21 to men. Also, to $s = o$, there are Table B.13, Table B.14, and Table B.15 to women, and Table B.22, Table B.23, and Table B.24 to men.

At the end, Table B.25 presents results of adjusted probabilities to models with aggregate effects using OLS, both to men and women. There were adjusted parameters of model in period 2003-2009.

Table B.1: $\mathbb{P}[s = \mathrm{e}] = \bar{X}\beta$ to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.462***	0.148***	0.870***	0.277***	0.497***	0.161***
	(0.008)	(0.00260)	(0.003)	(0.00110)	(0.008)	(0.00264)
30 to 34 years old	0.508***	0.163***	1.194***	0.380***	0.517***	0.167***
	(0.008)	(0.00263)	(0.004)	(0.00113)	(0.008)	(0.00266)
35 to 44 years old	0.595***	0.191***	1.273***	0.405***	0.585***	0.189***
	(0.007)	(0.00209)	(0.003)	(0.000973)	(0.007)	(0.00213)
45 to 54 years old	0.776***	0.249***	0.933***	0.297***	0.737***	0.239***
	(0.007)	(0.00211)	(0.003)	(0.00101)	(0.007)	(0.00217)
55 years old and over	-0.099***	-0.0317***	0.018***	0.00584***	-0.182***	-0.0588***
	(0.007)	(0.00214)	(0.003)	(0.00107)	(0.007)	(0.00231)
Married	-0.499***	-0.160***	-0.501***	-0.159***	-0.501***	-0.162***
	(0.002)	(0.000667)	(0.002)	(0.000659)	(0.002)	(0.000674)
Head of Household	0.378***	0.121***	0.371***	0.118***	0.378***	0.122***
	(0.003)	(0.000815)	(0.003)	(0.000804)	(0.003)	(0.000822)
Schooling						
Completed Primary	0.241***	0.0774***	0.072***	0.0230***	0.143***	0.0464***
	(0.003)	(0.00103)	(0.007)	(0.00227)	(0.007)	(0.00236)
Some Secondary Education	0.161***	0.0516***	0.181***	0.0576***	0.174***	0.0562***
	(0.003)	(0.000939)	(0.006)	(0.00197)	(0.007)	(0.00212)
Secondary School Graduated	0.655***	0.210***	0.323***	0.103***	0.376***	0.122***
	(0.003)	(0.000802)	(0.005)	(0.00168)	(0.006)	(0.00183)
Some Higher Education	0.653***	0.210***	0.448***	0.143***	0.453***	0.147***
	(0.003)	(0.00104)	(0.007)	(0.00226)	(0.007)	(0.00242)
Completed Higher Education	1.128***	0.362***	0.783***	0.249***	0.850***	0.275***
	(0.004)	(0.00124)	(0.009)	(0.00285)	(0.009)	(0.00290)
HDR	3.311***	1.063***	3.579***	1.140***	2.941***	0.952***
	(0.009)	(0.00298)	(0.008)	(0.00230)	(0.014)	(0.00433)
HDR×Age						
HDR×24-29 years old	0.927***	0.298***			0.841***	0.272***
	(0.018)	(0.00567)			(0.018)	(0.00578)
HDR×30-34 years old	1.714***	0.550***			1.695***	0.549***
	(0.019)	(0.00620)			(0.019)	(0.00629)
HDR×35-44 years old	1.678***	0.539***			1.716***	0.556***
	(0.014)	(0.00469)			(0.015)	(0.00481)
HDR×45-54 years old	0.318***	0.102***			0.426***	0.138***
	(0.013)	(0.00420)			(0.014)	(0.00440)
HDR×55 and over	0.272***	0.0872***			0.468***	0.152***
	(0.012)	(0.00388)			(0.013)	(0.00435)
HDR×Schooling						
HDR×Primary			0.368***	0.117***	0.207***	0.0670***
			(0.015)	(0.00470)	(0.015)	(0.00489)
HDR×Some Secondary			-0.068***	-0.0215***	-0.045***	-0.0146***
			(0.012)	(0.00397)	(0.013)	(0.00432)
HDR×Secondary Graduated			0.761***	0.242***	0.651***	0.211***
			(0.011)	(0.00350)	(0.012)	(0.00384)
HDR×Some Higher Education			0.466***	0.148***	0.453***	0.147***
			(0.015)	(0.00471)	(0.016)	(0.00508)
HDR×Higher Education Grad.			0.811***	0.258***	0.644***	0.208***
			(0.019)	(0.00620)	(0.019)	(0.00630)
Constant	-2.668***		-2.792***		-2.505***	
	(0.006)		(0.005)		(0.007)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.404		0.401		0.405	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.2: $\mathbb{P}[s=u]=\bar{X}\beta$ to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.483***	0.0206***	0.203***	0.00856***	0.453***	0.0194***
	(0.008)	(0.000352)	(0.004)	(0.000190)	(0.008)	(0.000361)
30 to 34 years old	0.362***	0.0154***	0.020***	0.000835***	0.350***	0.0150***
	(0.009)	(0.000378)	(0.005)	(0.000209)	(0.009)	(0.000387)
35 to 44 years old	0.258***	0.0110***	-0.106***	-0.00446***	0.260***	0.0112***
	(0.007)	(0.000310)	(0.004)	(0.000185)	(0.007)	(0.000317)
45 to 54 years old	0.009	0.000368	-0.185***	-0.00779***	0.016**	0.000693**
	(0.008)	(0.000337)	(0.005)	(0.000211)	(0.008)	(0.000347)
55 years old and over	-1.077***	-0.0459***	-0.993***	-0.0419***	-1.063***	-0.0456***
	(0.009)	(0.000388)	(0.007)	(0.000251)	(0.009)	(0.000401)
Married	-0.326***	-0.0139***	-0.334***	-0.0141***	-0.327***	-0.0140***
	(0.003)	(0.000145)	(0.003)	(0.000144)	(0.003)	(0.000146)
Head of Household	0.064***	0.00271***	0.072***	0.00305***	0.067***	0.00289***
	(0.004)	(0.000179)	(0.004)	(0.000176)	(0.004)	(0.000181)
Schooling						
Completed Primary	0.100***	0.00426***	0.193***	0.00812***	0.128***	0.00550***
	(0.006)	(0.000263)	(0.010)	(0.000421)	(0.010)	(0.000433)
Some Secondary Education	0.035***	0.00150***	0.087***	0.00366***	0.079***	0.00339***
	(0.006)	(0.000234)	(0.009)	(0.000368)	(0.009)	(0.000386)
Secondary School Graduated	0.516***	0.0220***	0.655***	0.0276***	0.610***	0.0262***
	(0.005)	(0.000199)	(0.007)	(0.000326)	(0.007)	(0.000329)
Some Higher Education	0.420***	0.0179***	0.540***	0.0228***	0.508***	0.0218***
	(0.006)	(0.000237)	(0.009)	(0.000391)	(0.009)	(0.000402)
Completed Higher Education	0.632***	0.0269***	0.977***	0.0412***	0.911***	0.0391***
	(0.006)	(0.000271)	(0.011)	(0.000478)	(0.011)	(0.000486)
HDR	-0.967***	-0.0412***	-0.952***	-0.0401***	-0.635***	-0.0273***
	(0.012)	(0.000527)	(0.018)	(0.000743)	(0.024)	(0.00102)
HDR×Age						
HDR×24-29 years old	-0.921***	-0.0393***			-0.814***	-0.0349***
	(0.021)	(0.000912)			(0.022)	(0.000946)
HDR×30-34 years old	-1.181***	-0.0503***			-1.128***	-0.0484***
	(0.025)	(0.00108)			(0.026)	(0.00112)
HDR×35-44 years old	-1.294***	-0.0551***			-1.303***	-0.0559***
	(0.021)	(0.000901)			(0.022)	(0.000928)
HDR×45-54 years old	-0.624***	-0.0266***			-0.656***	-0.0282***
	(0.020)	(0.000860)			(0.021)	(0.000899)
HDR×55 and over	0.644***	0.0274***			0.566***	0.0243***
	(0.021)	(0.000928)			(0.023)	(0.00102)
HDR×Schooling						
HDR×Primary			-0.400***	-0.0169***	-0.130***	-0.00557***
			(0.030)	(0.00129)	(0.031)	(0.00132)
HDR×Some Secondary			-0.277***	-0.0117***	-0.196***	-0.00841***
			(0.026)	(0.00111)	(0.027)	(0.00117)
HDR×Secondary Graduated			-0.593***	-0.0250***	-0.379***	-0.0163***
			(0.021)	(0.000910)	(0.022)	(0.000961)
HDR×Some Higher Education			-0.556***	-0.0234***	-0.361***	-0.0155***
			(0.025)	(0.00107)	(0.027)	(0.00114)
HDR×Higher Education Grad.			-1.245***	-0.0525***	-0.895***	-0.0384***
			(0.029)	(0.00124)	(0.030)	(0.00129)
Constant	-1.551***		-1.512***		-1.632***	
	(0.008)		(0.008)		(0.009)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.152		0.146		0.153	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.3: $\mathbb{P}[s = o] = \bar{X}\beta$ to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.485***	-0.171***	-0.845***	-0.297***	-0.503***	-0.178***
	(0.007)	(0.00238)	(0.003)	(0.00115)	(0.007)	(0.00240)
30 to 34 years old	-0.398***	-0.140***	-1.064***	-0.373***	-0.402***	-0.142***
	(0.007)	(0.00246)	(0.003)	(0.00119)	(0.007)	(0.00248)
35 to 44 years old	-0.394***	-0.138***	-1.099***	-0.386***	-0.387***	-0.136***
	(0.006)	(0.00194)	(0.003)	(0.00100)	(0.006)	(0.00197)
45 to 54 years old	-0.427***	-0.150***	-0.778***	-0.273***	-0.401***	-0.142***
	(0.006)	(0.00201)	(0.003)	(0.00105)	(0.006)	(0.00206)
55 years old and over	0.648***	0.228***	0.179***	0.0629***	0.695***	0.245***
	(0.006)	(0.00203)	(0.003)	(0.00111)	(0.006)	(0.00218)
Married	0.582***	0.204***	0.577***	0.202***	0.584***	0.206***
	(0.002)	(0.000707)	(0.002)	(0.000702)	(0.002)	(0.000711)
Head of Household	-0.388***	-0.136***	-0.373***	-0.131***	-0.388***	-0.137***
	(0.002)	(0.000866)	(0.002)	(0.000858)	(0.002)	(0.000870)
Schooling						
Completed Primary	-0.230***	-0.0808***	-0.172***	-0.0604***	-0.138***	-0.0485***
	(0.003)	(0.00108)	(0.006)	(0.00225)	(0.007)	(0.00232)
Some Secondary Education	-0.138***	-0.0486***	-0.272***	-0.0954***	-0.131***	-0.0464***
	(0.003)	(0.000976)	(0.006)	(0.00194)	(0.006)	(0.00207)
Secondary School Graduated	-0.723***	-0.254***	-0.645***	-0.226***	-0.553***	-0.195***
	(0.002)	(0.000842)	(0.005)	(0.00165)	(0.005)	(0.00178)
Some Higher Education	-0.691***	-0.243***	-0.699***	-0.245***	-0.549***	-0.194***
	(0.003)	(0.00109)	(0.006)	(0.00220)	(0.007)	(0.00232)
Completed Higher Education	-1.251***	-0.440***	-1.100***	-0.386***	-1.070***	-0.378***
	(0.004)	(0.00136)	(0.008)	(0.00286)	(0.008)	(0.00291)
HDR	-2.381***	-0.836***	-3.309***	-1.161***	-2.125***	-0.750***
	(0.008)	(0.00281)	(0.007)	(0.00242)	(0.012)	(0.00424)
HDR×Age						
HDR×24-29 years old	-0.872***	-0.306***			-0.827***	-0.292***
	(0.015)	(0.00546)			(0.016)	(0.00554)
HDR×30-34 years old	-1.755***	-0.616***			-1.746***	-0.616***
	(0.017)	(0.00607)			(0.017)	(0.00613)
HDR×35-44 years old	-1.842***	-0.647***			-1.874***	-0.661***
	(0.013)	(0.00454)			(0.013)	(0.00464)
HDR×45-54 years old	-0.857***	-0.301***			-0.934***	-0.330***
	(0.012)	(0.00411)			(0.012)	(0.00430)
HDR×55 and over	-1.101***	-0.387***			-1.224***	-0.432***
	(0.011)	(0.00379)			(0.012)	(0.00424)
HDR×Schooling						
HDR×Primary			-0.139***	-0.0487***	-0.210***	-0.0741***
			(0.014)	(0.00476)	(0.014)	(0.00493)
HDR×Some Secondary			0.337***	0.118***	0.000	9.68e-05
			(0.011)	(0.00401)	(0.012)	(0.00435)
HDR×Secondary Graduated			-0.175***	-0.0614***	-0.416***	-0.147***
			(0.010)	(0.00349)	(0.011)	(0.00383)
HDR×Some Higher Education			0.029**	0.0101**	-0.340***	-0.120***
			(0.013)	(0.00466)	(0.014)	(0.00503)
HDR×Higher Education Grad.			-0.373***	-0.131***	-0.439***	-0.155***
			(0.018)	(0.00633)	(0.018)	(0.00644)
Constant	2.031***		2.407***		1.923***	
	(0.005)		(0.005)		(0.006)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.365		0.360		0.365	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.4: $\mathbb{P}[s = \text{e}] = \bar{X}\beta$ to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.740***	0.211***	0.967***	0.282***	0.745***	0.211***
	(0.008)	(0.00237)	(0.004)	(0.00109)	(0.008)	(0.00238)
30 to 34 years old	0.932***	0.265***	1.294***	0.378***	0.959***	0.272***
	(0.009)	(0.00255)	(0.004)	(0.00124)	(0.009)	(0.00255)
35 to 44 years old	0.870***	0.248***	1.416***	0.413***	0.909***	0.258***
	(0.007)	(0.00213)	(0.004)	(0.00105)	(0.007)	(0.00215)
45 to 54 years old	0.624***	0.178***	1.042***	0.304***	0.666***	0.189***
	(0.007)	(0.00210)	(0.004)	(0.00112)	(0.007)	(0.00212)
55 years old and over	-0.821***	-0.234***	-0.316***	-0.0923***	-0.763***	-0.216***
	(0.006)	(0.00170)	(0.004)	(0.00105)	(0.006)	(0.00182)
Married	1.008***	0.287***	0.990***	0.289***	1.009***	0.286***
	(0.003)	(0.000844)	(0.003)	(0.000842)	(0.003)	(0.000842)
Head of Household	0.249***	0.0709***	0.241***	0.0704***	0.249***	0.0707***
	(0.003)	(0.000863)	(0.003)	(0.000874)	(0.003)	(0.000859)
Schooling						
Completed Primary	0.158***	0.0449***	0.268***	0.0782***	0.164***	0.0466***
	(0.003)	(0.000983)	(0.007)	(0.00197)	(0.007)	(0.00198)
Some Secondary Education	-0.167***	-0.0476***	0.255***	0.0743***	0.075***	0.0214***
	(0.003)	(0.000864)	(0.006)	(0.00166)	(0.006)	(0.00175)
Secondary School Graduated	0.318***	0.0905***	0.330***	0.0962***	0.202***	0.0574***
	(0.003)	(0.000801)	(0.005)	(0.00157)	(0.006)	(0.00162)
Some Higher Education	-0.204***	-0.0580***	0.150***	0.0439***	-0.022***	-0.00620***
	(0.004)	(0.00105)	(0.007)	(0.00214)	(0.008)	(0.00219)
Completed Higher Education	0.114***	0.0324***	0.333***	0.0973***	0.280***	0.0795***
	(0.004)	(0.00126)	(0.009)	(0.00258)	(0.009)	(0.00256)
HDR	3.243***	0.923***	4.391***	1.281***	3.444***	0.977***
	(0.008)	(0.00262)	(0.009)	(0.00262)	(0.014)	(0.00416)
HDR×Age						
HDR×24-29 years old	0.478***	0.136***			0.483***	0.137***
	(0.019)	(0.00527)			(0.019)	(0.00536)
HDR×30-34 years old	0.843***	0.240***			0.790***	0.224***
	(0.023)	(0.00639)			(0.023)	(0.00639)
HDR×35-44 years old	1.526***	0.435***			1.430***	0.406***
	(0.020)	(0.00546)			(0.020)	(0.00552)
HDR×45-54 years old	1.005***	0.286***			0.904***	0.256***
	(0.018)	(0.00499)			(0.018)	(0.00506)
HDR×55 and over	1.200***	0.342***			1.061***	0.301***
	(0.012)	(0.00331)			(0.013)	(0.00378)
HDR×Schooling						
HDR×Primary			-0.307***	-0.0897***	-0.014	-0.00407
			(0.017)	(0.00491)	(0.018)	(0.00499)
HDR×Some Secondary			-1.131***	-0.330***	-0.645***	-0.183***
			(0.013)	(0.00387)	(0.015)	(0.00423)
HDR×Secondary Graduated			-0.068***	-0.0198***	0.313***	0.0888***
			(0.013)	(0.00382)	(0.014)	(0.00402)
HDR×Some Higher Education			-0.941***	-0.275***	-0.466***	-0.132***
			(0.016)	(0.00475)	(0.018)	(0.00503)
HDR×Higher Education Grad.			-0.611***	-0.178***	-0.459***	-0.130***
			(0.021)	(0.00610)	(0.022)	(0.00611)
Constant	-1.857***		-2.319***		-1.937***	
	(0.006)		(0.005)		(0.007)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.487		0.486		0.488	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.5: $\mathbb{P}[s = u] = \bar{X}\beta$ to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.805***	0.0461***	0.340***	0.0207***	0.832***	0.0477***
	(0.008)	(0.000495)	(0.004)	(0.000267)	(0.009)	(0.000505)
30 to 34 years old	0.804***	0.0460***	0.198***	0.0121***	0.809***	0.0463***
	(0.009)	(0.000525)	(0.005)	(0.000293)	(0.009)	(0.000530)
35 to 44 years old	0.769***	0.0440***	0.028***	0.00173***	0.763***	0.0437***
	(0.008)	(0.000429)	(0.004)	(0.000264)	(0.008)	(0.000435)
45 to 54 years old	0.576***	0.0330***	0.067***	0.00411***	0.571***	0.0327***
	(0.008)	(0.000441)	(0.005)	(0.000293)	(0.008)	(0.000449)
55 years old and over	-0.438***	-0.0251***	-0.483***	-0.0295***	-0.430***	-0.0246***
	(0.007)	(0.000416)	(0.005)	(0.000320)	(0.007)	(0.000436)
Married	-0.186***	-0.0107***	-0.172***	-0.0105***	-0.189***	-0.0109***
	(0.004)	(0.000212)	(0.004)	(0.000221)	(0.004)	(0.000213)
Head of Household	-0.322***	-0.0184***	-0.306***	-0.0187***	-0.322***	-0.0184***
	(0.004)	(0.000219)	(0.004)	(0.000229)	(0.004)	(0.000220)
Schooling						
Completed Primary	0.124***	0.00708***	0.325***	0.0198***	0.244***	0.0140***
	(0.005)	(0.000277)	(0.008)	(0.000484)	(0.008)	(0.000467)
Some Secondary Education	0.002	0.000130	0.026***	0.00160***	0.035***	0.00198***
	(0.004)	(0.000254)	(0.007)	(0.000424)	(0.007)	(0.000423)
Secondary School Graduated	0.328***	0.0188***	0.443***	0.0270***	0.391***	0.0224***
	(0.004)	(0.000225)	(0.006)	(0.000387)	(0.006)	(0.000372)
Some Higher Education	0.087***	0.00496***	-0.018**	-0.00109**	-0.060***	-0.00345***
	(0.005)	(0.000296)	(0.008)	(0.000508)	(0.009)	(0.000503)
Completed Higher Education	0.282***	0.0161***	0.375***	0.0229***	0.235***	0.0135***
	(0.006)	(0.000333)	(0.010)	(0.000605)	(0.010)	(0.000582)
HDR	-1.380***	-0.0790***	-1.931***	-0.118***	-1.280***	-0.0733***
	(0.011)	(0.000654)	(0.014)	(0.000788)	(0.020)	(0.00113)
HDR×Age						
HDR×24-29 years old	-1.425***	-0.0815***			-1.508***	-0.0864***
	(0.021)	(0.00123)			(0.022)	(0.00127)
HDR×30-34 years old	-1.933***	-0.111***			-1.958***	-0.112***
	(0.025)	(0.00146)			(0.026)	(0.00147)
HDR×35-44 years old	-2.630***	-0.150***			-2.610***	-0.150***
	(0.023)	(0.00127)			(0.024)	(0.00130)
HDR×45-54 years old	-1.642***	-0.0940***			-1.625***	-0.0931***
	(0.022)	(0.00121)			(0.022)	(0.00124)
HDR×55 and over	0.444***	0.0254***			0.403***	0.0231***
	(0.017)	(0.000995)			(0.019)	(0.00110)
HDR×Schooling						
HDR×Primary			-0.810***	-0.0494***	-0.459***	-0.0263***
			(0.025)	(0.00152)	(0.026)	(0.00147)
HDR×Some Secondary			-0.190***	-0.0116***	-0.135***	-0.00771***
			(0.021)	(0.00129)	(0.023)	(0.00129)
HDR×Secondary Graduated			-0.473***	-0.0288***	-0.226***	-0.0129***
			(0.018)	(0.00111)	(0.019)	(0.00110)
HDR×Some Higher Education			0.216***	0.0132***	0.452***	0.0259***
			(0.023)	(0.00139)	(0.024)	(0.00140)
HDR×Higher Education Grad.			-0.453***	-0.0276***	0.120***	0.00688***
			(0.027)	(0.00164)	(0.027)	(0.00157)
Constant	-1.129***		-0.923***		-1.150***	
	(0.007)		(0.007)		(0.008)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.171		0.153		0.172	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.6: $\mathbb{P}[s=o]=\bar{X}\beta$ to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-1.161***	-0.257***	-1.200***	-0.264***	-1.166***	-0.257***
	(0.008)	(0.00185)	(0.004)	(0.000939)	(0.008)	(0.00187)
30 to 34 years old	-1.440***	-0.318***	-1.604***	-0.353***	-1.459***	-0.321***
	(0.010)	(0.00224)	(0.005)	(0.00116)	(0.010)	(0.00223)
35 to 44 years old	-1.239***	-0.274***	-1.595***	-0.351***	-1.270***	-0.280***
	(0.007)	(0.00177)	(0.004)	(0.000913)	(0.007)	(0.00179)
45 to 54 years old	-0.629***	-0.139***	-1.053***	-0.232***	-0.666***	-0.147***
	(0.007)	(0.00162)	(0.004)	(0.000933)	(0.007)	(0.00164)
55 years old and over	1.184***	0.262***	0.512***	0.113***	1.123***	0.247***
	(0.006)	(0.00126)	(0.004)	(0.000821)	(0.006)	(0.00133)
Married	-1.000***	-0.221***	-0.978***	-0.216***	-0.999***	-0.220***
	(0.003)	(0.000717)	(0.003)	(0.000679)	(0.003)	(0.000717)
Head of Household	-0.197***	-0.0435***	-0.208***	-0.0458***	-0.197***	-0.0435***
	(0.003)	(0.000735)	(0.003)	(0.000721)	(0.003)	(0.000732)
Schooling						
Completed Primary	-0.206***	-0.0456***	-0.475***	-0.105***	-0.271***	-0.0598***
	(0.004)	(0.000801)	(0.007)	(0.00145)	(0.007)	(0.00151)
Some Secondary Education	0.171***	0.0379***	-0.345***	-0.0760***	-0.052***	-0.0115***
	(0.003)	(0.000685)	(0.005)	(0.00119)	(0.006)	(0.00130)
Secondary School Graduated	-0.533***	-0.118***	-0.717***	-0.158***	-0.491***	-0.108***
	(0.003)	(0.000668)	(0.005)	(0.00117)	(0.006)	(0.00124)
Some Higher Education	0.185***	0.0410***	-0.220***	-0.0485***	0.091***	0.0201***
	(0.004)	(0.000828)	(0.007)	(0.00152)	(0.007)	(0.00162)
Completed Higher Education	-0.252***	-0.0557***	-0.543***	-0.120***	-0.405***	-0.0893***
	(0.005)	(0.00107)	(0.009)	(0.00196)	(0.009)	(0.00202)
HDR	-2.350***	-0.520***	-3.920***	-0.864***	-2.593***	-0.571***
	(0.007)	(0.00186)	(0.008)	(0.00200)	(0.013)	(0.00306)
HDR×Age						
HDR×24-29 years old	-0.018	-0.00393			-0.017	-0.00380
	(0.018)	(0.00405)			(0.019)	(0.00414)
HDR×30-34 years old	-0.283***	-0.0625***			-0.247***	-0.0545***
	(0.025)	(0.00553)			(0.025)	(0.00552)
HDR×35-44 years old	-0.910***	-0.201***			-0.829***	-0.183***
	(0.021)	(0.00454)			(0.021)	(0.00457)
HDR×45-54 years old	-1.126***	-0.249***			-1.034***	-0.228***
	(0.018)	(0.00389)			(0.018)	(0.00394)
HDR×55 and over	-1.794***	-0.397***			-1.640***	-0.361***
	(0.011)	(0.00241)			(0.012)	(0.00275)
HDR×Schooling						
HDR×Primary			0.786***	0.173***	0.186***	0.0411***
			(0.016)	(0.00357)	(0.017)	(0.00376)
HDR×Some Secondary			1.456***	0.321***	0.612***	0.135***
			(0.012)	(0.00277)	(0.014)	(0.00315)
HDR×Secondary Graduated			0.595***	0.131***	-0.115***	-0.0253***
			(0.013)	(0.00283)	(0.014)	(0.00306)
HDR×Some Higher Education			1.165***	0.257***	0.261***	0.0574***
			(0.015)	(0.00340)	(0.017)	(0.00375)
HDR×Higher Education Grad.			0.844***	0.186***	0.435***	0.0958***
			(0.021)	(0.00458)	(0.022)	(0.00478)
Constant	1.320***		1.897***		1.409***	
	(0.005)		(0.005)		(0.006)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.490		0.486		0.491	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.7: $\mathbb{P}[s = \mathrm{e}] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Age to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.458***	0.147***	0.464***	0.149***	0.304***	0.0975***
	(0.013)	(0.00411)	(0.008)	(0.00260)	(0.016)	(0.00501)
30 to 34 years old	0.503***	0.161***	0.510***	0.164***	0.340***	0.109***
	(0.013)	(0.00427)	(0.008)	(0.00263)	(0.016)	(0.00524)
35 to 44 years old	0.588***	0.189***	0.595***	0.191***	0.433***	0.139***
	(0.012)	(0.00386)	(0.007)	(0.00209)	(0.015)	(0.00485)
45 to 54 years old	0.768***	0.247***	0.774***	0.248***	0.628***	0.202***
	(0.011)	(0.00363)	(0.007)	(0.00211)	(0.014)	(0.00449)
55 years old and over	-0.103***	-0.0331***	-0.104***	-0.0333***	-0.082***	-0.0262***
	(0.007)	(0.00217)	(0.007)	(0.00214)	(0.007)	(0.00219)
Married	-0.499***	-0.160***	-0.499***	-0.160***	-0.498***	-0.160***
	(0.002)	(0.000667)	(0.002)	(0.000667)	(0.002)	(0.000667)
Head of Household	0.375***	0.120***	0.375***	0.120***	0.377***	0.121***
	(0.003)	(0.000815)	(0.003)	(0.000815)	(0.003)	(0.000815)
Schooling						
Completed Primary	0.236***	0.0756***	0.255***	0.0819***	0.240***	0.0772***
	(0.003)	(0.00103)	(0.007)	(0.00236)	(0.010)	(0.00308)
Some Secondary Ed.	0.153***	0.0490***	0.161***	0.0515***	0.160***	0.0513***
	(0.003)	(0.000942)	(0.004)	(0.00128)	(0.005)	(0.00152)
Secondary School Grad.	0.645***	0.207***	0.686***	0.220***	0.655***	0.210***
	(0.003)	(0.000805)	(0.014)	(0.00443)	(0.019)	(0.00602)
Some HE	0.641***	0.206***	0.686***	0.220***	0.653***	0.210***
	(0.003)	(0.00104)	(0.015)	(0.00490)	(0.021)	(0.00665)
HE Grad.	1.115***	0.358***	1.196***	0.384***	1.130***	0.363***
	(0.004)	(0.00125)	(0.027)	(0.00879)	(0.037)	(0.0120)
HDR	3.298***	1.058***	3.298***	1.059***	3.309***	1.062***
	(0.009)	(0.00298)	(0.009)	(0.00298)	(0.009)	(0.00298)
HDR×Age						
HDR×24-29 years old	0.926***	0.297***	0.926***	0.297***	0.924***	0.297***
	(0.018)	(0.00567)	(0.018)	(0.00567)	(0.018)	(0.00567)
HDR×30-34 years old	1.714***	0.550***	1.715***	0.550***	1.711***	0.549***
	(0.019)	(0.00620)	(0.019)	(0.00620)	(0.019)	(0.00620)
HDR×35-44 years old	1.682***	0.540***	1.682***	0.540***	1.675***	0.538***
	(0.014)	(0.00469)	(0.014)	(0.00469)	(0.014)	(0.00469)
HDR×45-54 years old	0.323***	0.104***	0.323***	0.104***	0.318***	0.102***
	(0.013)	(0.00420)	(0.013)	(0.00420)	(0.013)	(0.00420)
HDR×55 and over	0.278***	0.0892***	0.278***	0.0892***	0.271***	0.0871***
	(0.012)	(0.00388)	(0.012)	(0.00388)	(0.012)	(0.00388)
$\bar{Z}_a^{s=\mathrm{e}}$	0.008	0.00257			0.205***	0.0658***
	(0.013)	(0.00409)			(0.017)	(0.00552)
$\bar{Z}_r^{s=\mathrm{e}}$	0.288***	0.0923***	0.299***	0.0959***		
	(0.008)	(0.00255)	(0.008)	(0.00251)		
$\bar{Z}_s^{s=\mathrm{e}}$			-0.056***	-0.0181***	-0.003	-0.000945
			(0.019)	(0.00607)	(0.026)	(0.00832)
Constant	-2.522***		-2.575***		-2.490***	
	(0.012)		(0.017)		(0.017)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.404		0.404		0.404	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.8: $\mathbb{P}[s = \text{e}] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Schooling to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.842***	0.268***	0.872***	0.278***	0.686***	0.219***
	(0.010)	(0.00333)	(0.003)	(0.00110)	(0.014)	(0.00439)
30 to 34 years old	1.165***	0.371***	1.196***	0.381***	1.000***	0.318***
	(0.011)	(0.00351)	(0.004)	(0.00113)	(0.015)	(0.00465)
35 to 44 years old	1.244***	0.396***	1.274***	0.406***	1.085***	0.346***
	(0.011)	(0.00335)	(0.003)	(0.000973)	(0.014)	(0.00444)
45 to 54 years old	0.905***	0.288***	0.933***	0.297***	0.762***	0.243***
	(0.010)	(0.00308)	(0.003)	(0.00101)	(0.013)	(0.00406)
55 years old and over	0.020***	0.00622***	0.016***	0.00521***	0.038***	0.0121***
	(0.004)	(0.00112)	(0.003)	(0.00107)	(0.004)	(0.00117)
Married	-0.501***	-0.159***	-0.501***	-0.159***	-0.500***	-0.159***
	(0.002)	(0.000658)	(0.002)	(0.000658)	(0.002)	(0.000659)
Head of Household	0.368***	0.117***	0.368***	0.117***	0.370***	0.118***
	(0.003)	(0.000804)	(0.003)	(0.000804)	(0.003)	(0.000805)
Schooling						
Completed Primary	0.068***	0.0215***	0.078***	0.0247***	0.076***	0.0242***
	(0.007)	(0.00227)	(0.010)	(0.00308)	(0.011)	(0.00365)
Some Secondary Ed.	0.173***	0.0550***	0.177***	0.0562***	0.182***	0.0579***
	(0.006)	(0.00197)	(0.007)	(0.00214)	(0.007)	(0.00228)
Secondary School Grad.	0.313***	0.0997***	0.334***	0.106***	0.331***	0.106***
	(0.005)	(0.00168)	(0.015)	(0.00462)	(0.019)	(0.00617)
Some HE	0.437***	0.139***	0.459***	0.146***	0.458***	0.146***
	(0.007)	(0.00226)	(0.017)	(0.00526)	(0.022)	(0.00693)
HE Grad.	0.768***	0.244***	0.809***	0.258***	0.802***	0.255***
	(0.009)	(0.00285)	(0.029)	(0.00908)	(0.038)	(0.0122)
HDR	3.570***	1.136***	3.570***	1.136***	3.576***	1.139***
	(0.008)	(0.00230)	(0.008)	(0.00230)	(0.008)	(0.00230)
HDR×Schooling						
HDR×Primary	0.366***	0.116***	0.366***	0.117***	0.368***	0.117***
	(0.015)	(0.00470)	(0.015)	(0.00470)	(0.015)	(0.00470)
HDR×Some Secondary	-0.068***	-0.0217***	-0.068***	-0.0216***	-0.069***	-0.0218***
	(0.012)	(0.00397)	(0.012)	(0.00397)	(0.012)	(0.00397)
HDR×Secondary Grad.	0.760***	0.242***	0.760***	0.242***	0.761***	0.242***
	(0.011)	(0.00350)	(0.011)	(0.00350)	(0.011)	(0.00350)
HDR×Some HE	0.465***	0.148***	0.465***	0.148***	0.466***	0.148***
	(0.015)	(0.00471)	(0.015)	(0.00471)	(0.015)	(0.00471)
HDR×HE Grad.	0.814***	0.259***	0.814***	0.259***	0.812***	0.258***
	(0.019)	(0.00619)	(0.019)	(0.00619)	(0.019)	(0.00620)
$\bar{Z}_a^{s=\text{e}}$	0.038***	0.0120***			0.235***	0.0749***
	(0.013)	(0.00401)			(0.017)	(0.00544)
$\bar{Z}_r^{s=\text{e}}$	0.281***	0.0894***	0.296***	0.0942***		
	(0.008)	(0.00253)	(0.008)	(0.00249)		
$\bar{Z}_s^{s=\text{e}}$			-0.029	-0.00917	-0.016	-0.00506
			(0.019)	(0.00602)	(0.026)	(0.00830)
Constant	-2.625***		-2.677***		-2.598***	
	(0.011)		(0.017)		(0.017)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.401		0.401		0.401	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.9: $\mathbb{P}[s = \mathrm{e}] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with two interaction terms to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.490***	0.159***	0.499***	0.161***	0.340***	0.110***
	(0.013)	(0.00415)	(0.008)	(0.00264)	(0.016)	(0.00506)
30 to 34 years old	0.509***	0.165***	0.518***	0.168***	0.350***	0.113***
	(0.013)	(0.00431)	(0.008)	(0.00266)	(0.016)	(0.00529)
35 to 44 years old	0.576***	0.186***	0.585***	0.189***	0.424***	0.137***
	(0.012)	(0.00390)	(0.007)	(0.00213)	(0.015)	(0.00490)
45 to 54 years old	0.727***	0.235***	0.735***	0.238***	0.590***	0.191***
	(0.011)	(0.00369)	(0.007)	(0.00217)	(0.014)	(0.00455)
55 years old and over	-0.186***	-0.0602***	-0.187***	-0.0605***	-0.165***	-0.0533***
	(0.007)	(0.00234)	(0.007)	(0.00231)	(0.007)	(0.00236)
Married	-0.502***	-0.162***	-0.502***	-0.162***	-0.501***	-0.162***
	(0.002)	(0.000674)	(0.002)	(0.000674)	(0.002)	(0.000674)
Head of Household	0.374***	0.121***	0.375***	0.121***	0.376***	0.122***
	(0.003)	(0.000822)	(0.003)	(0.000822)	(0.003)	(0.000822)
Schooling						
Completed Primary	0.138***	0.0447***	0.155***	0.0501***	0.139***	0.0451***
	(0.007)	(0.00236)	(0.010)	(0.00318)	(0.012)	(0.00374)
Some Secondary Ed.	0.164***	0.0532***	0.171***	0.0554***	0.171***	0.0554***
	(0.007)	(0.00212)	(0.007)	(0.00228)	(0.007)	(0.00242)
Secondary School Grad.	0.366***	0.118***	0.400***	0.130***	0.368***	0.119***
	(0.006)	(0.00183)	(0.015)	(0.00475)	(0.019)	(0.00628)
Some HE	0.441***	0.143***	0.479***	0.155***	0.445***	0.144***
	(0.008)	(0.00242)	(0.017)	(0.00541)	(0.022)	(0.00706)
HE Grad.	0.834***	0.270***	0.903***	0.292***	0.836***	0.270***
	(0.009)	(0.00291)	(0.029)	(0.00925)	(0.038)	(0.0124)
HDR	2.927***	0.947***	2.927***	0.947***	2.940***	0.952***
	(0.014)	(0.00433)	(0.014)	(0.00433)	(0.014)	(0.00433)
HDR×Age						
HDR×24-29 years old	0.840***	0.272***	0.840***	0.272***	0.838***	0.271***
	(0.018)	(0.00577)	(0.018)	(0.00577)	(0.018)	(0.00578)
HDR×30-34 years old	1.696***	0.549***	1.696***	0.549***	1.692***	0.548***
	(0.019)	(0.00629)	(0.019)	(0.00629)	(0.019)	(0.00629)
HDR×35-44 years old	1.720***	0.556***	1.720***	0.557***	1.713***	0.555***
	(0.015)	(0.00481)	(0.015)	(0.00481)	(0.015)	(0.00481)
HDR×45-54 years old	0.431***	0.139***	0.431***	0.139***	0.426***	0.138***
	(0.014)	(0.00440)	(0.014)	(0.00440)	(0.014)	(0.00440)
HDR×55 and over	0.475***	0.154***	0.475***	0.154***	0.468***	0.151***
	(0.013)	(0.00435)	(0.013)	(0.00435)	(0.013)	(0.00435)
HDR×Schooling						
HDR×Primary	0.206***	0.0668***	0.207***	0.0669***	0.207***	0.0669***
	(0.015)	(0.00489)	(0.015)	(0.00489)	(0.015)	(0.00489)
HDR×Some Secondary	-0.043***	-0.0140***	-0.043***	-0.0138***	-0.046***	-0.0148***
	(0.013)	(0.00432)	(0.013)	(0.00432)	(0.013)	(0.00432)
HDR×Secondary Grad.	0.653***	0.211***	0.653***	0.211***	0.652***	0.211***
	(0.012)	(0.00384)	(0.012)	(0.00384)	(0.012)	(0.00384)
HDR×Some HE	0.455***	0.147***	0.455***	0.147***	0.453***	0.147***
	(0.016)	(0.00507)	(0.016)	(0.00507)	(0.016)	(0.00508)
HDR×HE Grad.	0.649***	0.210***	0.649***	0.210***	0.645***	0.209***
	(0.019)	(0.00630)	(0.019)	(0.00630)	(0.019)	(0.00630)
$\bar{Z}_a^{s=\mathrm{e}}$	0.011	0.00356			0.203***	0.0657***
	(0.013)	(0.00413)			(0.017)	(0.00556)
$\bar{Z}_r^{s=\mathrm{e}}$	0.288***	0.0933***	0.299***	0.0967***		
	(0.008)	(0.00258)	(0.008)	(0.00253)		
$\bar{Z}_s^{s=\mathrm{e}}$			-0.048**	-0.0156**	0.008	0.00243
			(0.019)	(0.00613)	(0.026)	(0.00840)
Constant	-2.356***		-2.404***		-2.319***	
	(0.013)		(0.018)		(0.017)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.405		0.405		0.405	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.10: $\mathbb{P}[s = u] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Age to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.452***	0.0186***	0.488***	0.0202***	0.409***	0.0173***
	(0.009)	(0.000364)	(0.008)	(0.000344)	(0.009)	(0.000379)
30 to 34 years old	0.389***	0.0160***	0.365***	0.0151***	0.410***	0.0174***
	(0.009)	(0.000377)	(0.009)	(0.000369)	(0.009)	(0.000388)
35 to 44 years old	0.324***	0.0134***	0.258***	0.0106***	0.392***	0.0166***
	(0.009)	(0.000372)	(0.007)	(0.000302)	(0.010)	(0.000407)
45 to 54 years old	0.117***	0.00482***	0.003	0.000125	0.238***	0.0101***
	(0.012)	(0.000498)	(0.008)	(0.000329)	(0.013)	(0.000566)
55 years old and over	-0.799***	-0.0330***	-1.094***	-0.0452***	-0.480***	-0.0203***
	(0.025)	(0.00105)	(0.009)	(0.000380)	(0.030)	(0.00125)
Married	-0.325***	-0.0134***	-0.325***	-0.0134***	-0.325***	-0.0138***
	(0.003)	(0.000141)	(0.003)	(0.000141)	(0.003)	(0.000144)
Head of Household	0.056***	0.00232***	0.057***	0.00235***	0.061***	0.00256***
	(0.004)	(0.000174)	(0.004)	(0.000174)	(0.004)	(0.000178)
Schooling						
Completed Primary	0.089***	0.00366***	-0.001	-4.10e-05	-0.036***	-0.00153***
	(0.006)	(0.000257)	(0.011)	(0.000441)	(0.012)	(0.000511)
Some Secondary Ed.	0.015***	0.000602***	-0.086***	-0.00356***	-0.115***	-0.00488***
	(0.006)	(0.000228)	(0.011)	(0.000460)	(0.013)	(0.000542)
Secondary School Grad.	0.492***	0.0203***	0.285***	0.0118***	0.210***	0.00888***
	(0.005)	(0.000195)	(0.020)	(0.000838)	(0.024)	(0.00102)
Some HE	0.390***	0.0161***	0.197***	0.00815***	0.135***	0.00569***
	(0.006)	(0.000231)	(0.019)	(0.000794)	(0.023)	(0.000962)
HE Grad.	0.599***	0.0247***	0.406***	0.0168***	0.346***	0.0146***
	(0.006)	(0.000264)	(0.019)	(0.000804)	(0.023)	(0.000971)
HDR	-0.973***	-0.0402***	-0.973***	-0.0402***	-0.966***	-0.0409***
	(0.012)	(0.000514)	(0.012)	(0.000515)	(0.012)	(0.000524)
HDR×Age						
HDR×24-29 years old	-0.927***	-0.0383***	-0.926***	-0.0382***	-0.924***	-0.0391***
	(0.021)	(0.000888)	(0.021)	(0.000889)	(0.021)	(0.000907)
HDR×30-34 years old	-1.189***	-0.0491***	-1.186***	-0.0490***	-1.186***	-0.0502***
	(0.025)	(0.00106)	(0.025)	(0.00106)	(0.025)	(0.00108)
HDR×35-44 years old	-1.302***	-0.0537***	-1.299***	-0.0537***	-1.300***	-0.0550***
	(0.021)	(0.000878)	(0.021)	(0.000879)	(0.021)	(0.000895)
HDR×45-54 years old	-0.628***	-0.0259***	-0.626***	-0.0259***	-0.629***	-0.0266***
	(0.021)	(0.000840)	(0.021)	(0.000840)	(0.020)	(0.000855)
HDR×55 and over	0.647***	0.0267***	0.649***	0.0268***	0.641***	0.0271***
	(0.022)	(0.000908)	(0.022)	(0.000908)	(0.021)	(0.000923)
$\bar{Z}_a^{s=u}$	0.281***	0.0116***			0.573***	0.0242***
	(0.023)	(0.000932)			(0.027)	(0.00114)
$\bar{Z}_r^{s=u}$	0.885***	0.0365***	0.897***	0.0370***		
	(0.012)	(0.000504)	(0.012)	(0.000500)		
$\bar{Z}_s^{s=u}$			0.279***	0.0115***	0.404***	0.0171***
			(0.027)	(0.00110)	(0.032)	(0.00135)
Constant	0.567***		0.791***		0.327***	
	(0.037)		(0.061)		(0.060)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.158		0.158		0.153	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.11: $\mathbb{P}[s=u] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Schooling to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.173***	0.00707***	0.208***	0.00849***	0.131***	0.00549***
	(0.005)	(0.000221)	(0.004)	(0.000186)	(0.006)	(0.000239)
30 to 34 years old	0.043***	0.00175***	0.022***	0.000887***	0.065***	0.00270***
	(0.005)	(0.000216)	(0.005)	(0.000204)	(0.005)	(0.000226)
35 to 44 years old	-0.046***	-0.00190***	-0.108***	-0.00440***	0.022***	0.000904***
	(0.007)	(0.000277)	(0.004)	(0.000180)	(0.008)	(0.000317)
45 to 54 years old	-0.085***	-0.00348***	-0.191***	-0.00782***	0.034***	0.00143***
	(0.010)	(0.000419)	(0.005)	(0.000206)	(0.012)	(0.000493)
55 years old and over	-0.731***	-0.0299***	-1.008***	-0.0412***	-0.419***	-0.0175***
	(0.024)	(0.000989)	(0.007)	(0.000246)	(0.029)	(0.00120)
Married	-0.333***	-0.0136***	-0.333***	-0.0136***	-0.334***	-0.0140***
	(0.003)	(0.000140)	(0.003)	(0.000140)	(0.003)	(0.000143)
Head of Household	0.065***	0.00265***	0.066***	0.00268***	0.069***	0.00291***
	(0.004)	(0.000171)	(0.004)	(0.000171)	(0.004)	(0.000175)
Schooling						
Completed Primary	0.182***	0.00743***	0.094***	0.00386***	0.056***	0.00234***
	(0.010)	(0.000411)	(0.013)	(0.000540)	(0.014)	(0.000599)
Some Secondary Ed.	0.067***	0.00274***	-0.031**	-0.00125**	-0.064***	-0.00269***
	(0.009)	(0.000359)	(0.013)	(0.000529)	(0.014)	(0.000602)
Secondary School Grad.	0.632***	0.0258***	0.430***	0.0176***	0.346***	0.0145***
	(0.007)	(0.000318)	(0.021)	(0.000862)	(0.025)	(0.00104)
Some HE	0.508***	0.0208***	0.320***	0.0131***	0.250***	0.0105***
	(0.009)	(0.000380)	(0.020)	(0.000837)	(0.024)	(0.000996)
HE Grad.	0.945***	0.0386***	0.757***	0.0310***	0.688***	0.0288***
	(0.011)	(0.000465)	(0.021)	(0.000878)	(0.025)	(0.00103)
HDR	-0.959***	-0.0392***	-0.958***	-0.0392***	-0.952***	-0.0399***
	(0.018)	(0.000725)	(0.018)	(0.000725)	(0.018)	(0.000738)
HDR×Schooling						
HDR×Primary	-0.402***	-0.0164***	-0.403***	-0.0165***	-0.404***	-0.0169***
	(0.031)	(0.00126)	(0.031)	(0.00126)	(0.030)	(0.00128)
HDR×Some Secondary	-0.281***	-0.0115***	-0.283***	-0.0116***	-0.282***	-0.0118***
	(0.026)	(0.00108)	(0.026)	(0.00108)	(0.026)	(0.00110)
HDR×Secondary Grad.	-0.596***	-0.0244***	-0.595***	-0.0244***	-0.594***	-0.0249***
	(0.021)	(0.000889)	(0.021)	(0.000889)	(0.021)	(0.000905)
HDR×Some HE	-0.552***	-0.0226***	-0.552***	-0.0226***	-0.556***	-0.0233***
	(0.025)	(0.00105)	(0.025)	(0.00105)	(0.025)	(0.00107)
HDR×HE Grad.	-1.250***	-0.0511***	-1.249***	-0.0511***	-1.247***	-0.0522***
	(0.029)	(0.00121)	(0.029)	(0.00121)	(0.029)	(0.00124)
$\bar{Z}_a^{s=u}$	0.264***	0.0108***			0.552***	0.0231***
	(0.022)	(0.000916)			(0.027)	(0.00112)
$\bar{Z}_r^{s=u}$	0.884***	0.0361***	0.894***	0.0366***		
	(0.012)	(0.000499)	(0.012)	(0.000494)		
$\bar{Z}_s^{s=u}$			0.272***	0.0111***	0.409***	0.0171***
			(0.026)	(0.00108)	(0.032)	(0.00133)
Constant	0.578***		0.810***		0.345***	
	(0.037)		(0.061)		(0.060)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.152		0.152		0.147	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.12: $\mathbb{P}[s=u] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with two interaction terms to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.422***	0.0175***	0.459***	0.0191***	0.378***	0.0161***
	(0.009)	(0.000372)	(0.008)	(0.000352)	(0.009)	(0.000387)
30 to 34 years old	0.376***	0.0157***	0.353***	0.0147***	0.399***	0.0170***
	(0.009)	(0.000385)	(0.009)	(0.000377)	(0.009)	(0.000396)
35 to 44 years old	0.326***	0.0135***	0.259***	0.0108***	0.395***	0.0169***
	(0.009)	(0.000379)	(0.007)	(0.000309)	(0.010)	(0.000414)
45 to 54 years old	0.125***	0.00521***	0.011	0.000446	0.247***	0.0105***
	(0.012)	(0.000506)	(0.008)	(0.000338)	(0.014)	(0.000575)
55 years old and over	-0.781***	-0.0325***	-1.079***	-0.0449***	-0.460***	-0.0196***
	(0.025)	(0.00106)	(0.010)	(0.000392)	(0.030)	(0.00126)
Married	-0.326***	-0.0136***	-0.326***	-0.0136***	-0.326***	-0.0139***
	(0.003)	(0.000142)	(0.003)	(0.000142)	(0.003)	(0.000145)
Head of Household	0.060***	0.00249***	0.061***	0.00252***	0.064***	0.00274***
	(0.004)	(0.000176)	(0.004)	(0.000176)	(0.004)	(0.000180)
Schooling						
Completed Primary	0.117***	0.00485***	0.029**	0.00119**	-0.004	-0.000183
	(0.010)	(0.000422)	(0.013)	(0.000555)	(0.014)	(0.000615)
Some Secondary Ed.	0.059***	0.00244***	-0.040***	-0.00167***	-0.067***	-0.00284***
	(0.009)	(0.000377)	(0.013)	(0.000548)	(0.015)	(0.000623)
Secondary School Grad.	0.585***	0.0244***	0.381***	0.0159***	0.310***	0.0132***
	(0.008)	(0.000321)	(0.021)	(0.000881)	(0.025)	(0.00106)
Some HE	0.476***	0.0198***	0.286***	0.0119***	0.228***	0.00973***
	(0.009)	(0.000392)	(0.021)	(0.000858)	(0.024)	(0.00102)
HE Grad.	0.878***	0.0365***	0.688***	0.0286***	0.630***	0.0269***
	(0.011)	(0.000474)	(0.022)	(0.000899)	(0.025)	(0.00106)
HDR	-0.641***	-0.0267***	-0.642***	-0.0267***	-0.631***	-0.0269***
	(0.024)	(0.000993)	(0.024)	(0.000994)	(0.024)	(0.00101)
HDR×Age						
HDR×24-29 years old	-0.819***	-0.0341***	-0.819***	-0.0341***	-0.818***	-0.0349***
	(0.022)	(0.000921)	(0.022)	(0.000921)	(0.022)	(0.000940)
HDR×30-34 years old	-1.134***	-0.0472***	-1.131***	-0.0471***	-1.133***	-0.0483***
	(0.026)	(0.00109)	(0.026)	(0.00109)	(0.026)	(0.00111)
HDR×35-44 years old	-1.310***	-0.0545***	-1.307***	-0.0544***	-1.310***	-0.0558***
	(0.022)	(0.000905)	(0.022)	(0.000905)	(0.022)	(0.000922)
HDR×45-54 years old	-0.662***	-0.0276***	-0.660***	-0.0275***	-0.663***	-0.0282***
	(0.021)	(0.000878)	(0.021)	(0.000878)	(0.021)	(0.000894)
HDR×55 and over	0.567***	0.0236***	0.569***	0.0237***	0.561***	0.0239***
	(0.024)	(0.000994)	(0.024)	(0.000994)	(0.023)	(0.00101)
HDR×Schooling						
HDR×Primary	-0.129***	-0.00536***	-0.130***	-0.00540***	-0.134***	-0.00572***
	(0.031)	(0.00129)	(0.031)	(0.00129)	(0.031)	(0.00131)
HDR×Some Secondary	-0.198***	-0.00822***	-0.199***	-0.00827***	-0.204***	-0.00868***
	(0.028)	(0.00115)	(0.028)	(0.00115)	(0.027)	(0.00117)
HDR×Secondary Grad.	-0.380***	-0.0158***	-0.378***	-0.0157***	-0.381***	-0.0162***
	(0.022)	(0.000939)	(0.022)	(0.000940)	(0.022)	(0.000955)
HDR×Some HE	-0.356***	-0.0148***	-0.355***	-0.0148***	-0.362***	-0.0154***
	(0.027)	(0.00111)	(0.027)	(0.00111)	(0.027)	(0.00113)
HDR×HE Grad.	-0.895***	-0.0372***	-0.894***	-0.0372***	-0.897***	-0.0382***
	(0.030)	(0.00126)	(0.030)	(0.00126)	(0.030)	(0.00128)
$\bar{Z}_a^{s=u}$	0.283***	0.0118***			0.578***	0.0246***
	(0.023)	(0.000940)			(0.027)	(0.00115)
$\bar{Z}_r^{s=u}$	0.884***	0.0368***	0.896***	0.0373***		
	(0.012)	(0.000508)	(0.012)	(0.000503)		
$\bar{Z}_s^{s=u}$			0.275***	0.0114***	0.396***	0.0169***
			(0.027)	(0.00110)	(0.032)	(0.00136)
Constant	0.488***		0.701***		0.235***	
	(0.038)		(0.061)		(0.061)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.159		0.159		0.154	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.13: $\mathbb{P}[s=o] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Age to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.386***	-0.136***	-0.488***	-0.171***	-0.245***	-0.0861***
	(0.011)	(0.00385)	(0.007)	(0.00238)	(0.013)	(0.00474)
30 to 34 years old	-0.300***	-0.105***	-0.399***	-0.140***	-0.163***	-0.0572***
	(0.011)	(0.00385)	(0.007)	(0.00246)	(0.013)	(0.00470)
35 to 44 years old	-0.303***	-0.106***	-0.393***	-0.138***	-0.180***	-0.0632***
	(0.009)	(0.00332)	(0.006)	(0.00194)	(0.012)	(0.00415)
45 to 54 years old	-0.346***	-0.121***	-0.423***	-0.148***	-0.242***	-0.0851***
	(0.009)	(0.00306)	(0.006)	(0.00201)	(0.011)	(0.00370)
55 years old and over	0.622***	0.218***	0.658***	0.231***	0.566***	0.199***
	(0.007)	(0.00230)	(0.006)	(0.00203)	(0.007)	(0.00250)
Married	0.582***	0.204***	0.582***	0.204***	0.582***	0.204***
	(0.002)	(0.000707)	(0.002)	(0.000707)	(0.002)	(0.000707)
Head of Household	-0.382***	-0.134***	-0.382***	-0.134***	-0.385***	-0.135***
	(0.002)	(0.000867)	(0.002)	(0.000867)	(0.002)	(0.000867)
Schooling						
Completed Primary	-0.222***	-0.0778***	-0.176***	-0.0620***	-0.181***	-0.0637***
	(0.003)	(0.00108)	(0.007)	(0.00256)	(0.010)	(0.00335)
Some Secondary Ed.	-0.125***	-0.0440***	-0.103***	-0.0361***	-0.113***	-0.0395***
	(0.003)	(0.000979)	(0.004)	(0.00151)	(0.005)	(0.00186)
Secondary School Grad.	-0.708***	-0.249***	-0.609***	-0.214***	-0.619***	-0.217***
	(0.002)	(0.000846)	(0.015)	(0.00513)	(0.020)	(0.00700)
Some HE	-0.673***	-0.236***	-0.567***	-0.199***	-0.580***	-0.204***
	(0.003)	(0.00109)	(0.016)	(0.00552)	(0.021)	(0.00751)
HE Grad.	-1.231***	-0.432***	-1.044***	-0.366***	-1.059***	-0.372***
	(0.004)	(0.00137)	(0.027)	(0.00961)	(0.037)	(0.0131)
HDR	-2.364***	-0.830***	-2.362***	-0.830***	-2.377***	-0.835***
	(0.008)	(0.00281)	(0.008)	(0.00281)	(0.008)	(0.00281)
HDR×Age						
HDR×24-29 years old	-0.872***	-0.306***	-0.874***	-0.307***	-0.869***	-0.305***
	(0.015)	(0.00546)	(0.015)	(0.00546)	(0.015)	(0.00546)
HDR×30-34 years old	-1.757***	-0.617***	-1.759***	-0.618***	-1.752***	-0.615***
	(0.017)	(0.00607)	(0.017)	(0.00607)	(0.017)	(0.00607)
HDR×35-44 years old	-1.848***	-0.649***	-1.849***	-0.649***	-1.838***	-0.646***
	(0.013)	(0.00454)	(0.013)	(0.00454)	(0.013)	(0.00454)
HDR×45-54 years old	-0.864***	-0.303***	-0.865***	-0.304***	-0.856***	-0.301***
	(0.012)	(0.00411)	(0.012)	(0.00411)	(0.012)	(0.00411)
HDR×55 and over	-1.111***	-0.390***	-1.112***	-0.390***	-1.100***	-0.387***
	(0.011)	(0.00379)	(0.011)	(0.00379)	(0.011)	(0.00379)
$\bar{Z}_a^{s=o}$	0.132***	0.0463***			0.314***	0.110***
	(0.011)	(0.00394)			(0.015)	(0.00534)
$\bar{Z}_r^{s=o}$	0.368***	0.129***	0.384***	0.135***		
	(0.007)	(0.00255)	(0.007)	(0.00252)		
$\bar{Z}_s^{s=o}$			0.120***	0.0423***	0.120***	0.0423***
			(0.017)	(0.00613)	(0.024)	(0.00841)
Constant	1.793***		1.772***		1.708***	
	(0.009)		(0.015)		(0.016)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.366		0.366		0.365	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.14: $\mathbb{P}[s=o] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Schooling to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.724***	-0.254***	-0.848***	-0.298***	-0.581***	-0.204***
	(0.009)	(0.00323)	(0.003)	(0.00115)	(0.012)	(0.00426)
30 to 34 years old	-0.946***	-0.332***	-1.067***	-0.374***	-0.806***	-0.283***
	(0.009)	(0.00317)	(0.003)	(0.00119)	(0.012)	(0.00418)
35 to 44 years old	-0.990***	-0.347***	-1.101***	-0.386***	-0.862***	-0.303***
	(0.008)	(0.00287)	(0.003)	(0.00100)	(0.011)	(0.00380)
45 to 54 years old	-0.682***	-0.239***	-0.777***	-0.272***	-0.575***	-0.202***
	(0.007)	(0.00252)	(0.003)	(0.00105)	(0.009)	(0.00328)
55 years old and over	0.140***	0.0492***	0.183***	0.0643***	0.089***	0.0311***
	(0.004)	(0.00152)	(0.003)	(0.00111)	(0.005)	(0.00180)
Married	0.577***	0.202***	0.577***	0.202***	0.577***	0.202***
	(0.002)	(0.000702)	(0.002)	(0.000702)	(0.002)	(0.000703)
Head of Household	-0.367***	-0.129***	-0.367***	-0.129***	-0.370***	-0.130***
	(0.002)	(0.000858)	(0.002)	(0.000858)	(0.002)	(0.000858)
Schooling						
Completed Primary	-0.165***	-0.0578***	-0.111***	-0.0390***	-0.135***	-0.0475***
	(0.006)	(0.00225)	(0.009)	(0.00321)	(0.011)	(0.00388)
Some Secondary Ed.	-0.260***	-0.0913***	-0.234***	-0.0820***	-0.253***	-0.0887***
	(0.006)	(0.00195)	(0.006)	(0.00224)	(0.007)	(0.00249)
Secondary School Grad.	-0.631***	-0.221***	-0.512***	-0.180***	-0.566***	-0.199***
	(0.005)	(0.00165)	(0.015)	(0.00530)	(0.020)	(0.00715)
Some HE	-0.682***	-0.239***	-0.555***	-0.195***	-0.614***	-0.216***
	(0.006)	(0.00220)	(0.017)	(0.00582)	(0.022)	(0.00778)
HE Grad.	-1.078***	-0.378***	-0.855***	-0.300***	-0.954***	-0.335***
	(0.008)	(0.00286)	(0.028)	(0.00991)	(0.038)	(0.0134)
HDR	-3.299***	-1.157***	-3.299***	-1.157***	-3.305***	-1.160***
	(0.007)	(0.00242)	(0.007)	(0.00242)	(0.007)	(0.00242)
HDR×Schooling						
HDR×Primary	-0.136***	-0.0478***	-0.135***	-0.0475***	-0.138***	-0.0483***
	(0.014)	(0.00476)	(0.014)	(0.00476)	(0.014)	(0.00476)
HDR×Some Secondary	0.339***	0.119***	0.340***	0.119***	0.339***	0.119***
	(0.011)	(0.00401)	(0.011)	(0.00401)	(0.011)	(0.00401)
HDR×Secondary Grad.	-0.172***	-0.0603***	-0.171***	-0.0601***	-0.174***	-0.0612***
	(0.010)	(0.00349)	(0.010)	(0.00349)	(0.010)	(0.00349)
HDR×Some HE	0.031**	0.0107**	0.030**	0.0107**	0.029**	0.0101**
	(0.013)	(0.00466)	(0.013)	(0.00466)	(0.013)	(0.00466)
HDR×HE Grad.	-0.377***	-0.132***	-0.377***	-0.132***	-0.375***	-0.132***
	(0.018)	(0.00633)	(0.018)	(0.00633)	(0.018)	(0.00633)
$\bar{Z}_a^{s=o}$	0.161***	0.0563***			0.343***	0.120***
	(0.011)	(0.00390)			(0.015)	(0.00532)
$\bar{Z}_r^{s=o}$	0.348***	0.122***	0.369***	0.130***		
	(0.007)	(0.00254)	(0.007)	(0.00251)		
$\bar{Z}_s^{s=o}$			0.144***	0.0505***	0.090***	0.0316***
			(0.017)	(0.00610)	(0.024)	(0.00843)
Constant	2.160***		2.138***		2.090***	
	(0.009)		(0.015)		(0.016)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.361		0.361		0.360	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.15: $\mathbb{P}[s=o] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with two interaction terms to women

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.403***	-0.142***	-0.506***	-0.178***	-0.266***	-0.0938***
	(0.011)	(0.00388)	(0.007)	(0.00240)	(0.014)	(0.00477)
30 to 34 years old	-0.303***	-0.107***	-0.404***	-0.142***	-0.171***	-0.0602***
	(0.011)	(0.00387)	(0.007)	(0.00248)	(0.013)	(0.00473)
35 to 44 years old	-0.294***	-0.104***	-0.386***	-0.136***	-0.175***	-0.0617***
	(0.010)	(0.00335)	(0.006)	(0.00197)	(0.012)	(0.00417)
45 to 54 years old	-0.319***	-0.112***	-0.397***	-0.140***	-0.219***	-0.0771***
	(0.009)	(0.00310)	(0.006)	(0.00206)	(0.011)	(0.00374)
55 years old and over	0.668***	0.236***	0.705***	0.249***	0.614***	0.216***
	(0.007)	(0.00243)	(0.006)	(0.00218)	(0.007)	(0.00262)
Married	0.584***	0.206***	0.584***	0.206***	0.584***	0.206***
	(0.002)	(0.000711)	(0.002)	(0.000711)	(0.002)	(0.000711)
Head of Household	-0.382***	-0.135***	-0.383***	-0.135***	-0.385***	-0.136***
	(0.002)	(0.000871)	(0.002)	(0.000870)	(0.002)	(0.000871)
Schooling						
Completed Primary	-0.129***	-0.0455***	-0.081***	-0.0286***	-0.085***	-0.0300***
	(0.007)	(0.00232)	(0.009)	(0.00328)	(0.011)	(0.00393)
Some Secondary Ed.	-0.118***	-0.0415***	-0.094***	-0.0331***	-0.105***	-0.0369***
	(0.006)	(0.00207)	(0.007)	(0.00236)	(0.007)	(0.00259)
Secondary School Grad.	-0.537***	-0.189***	-0.432***	-0.152***	-0.440***	-0.155***
	(0.005)	(0.00178)	(0.015)	(0.00539)	(0.020)	(0.00720)
Some HE	-0.530***	-0.187***	-0.417***	-0.147***	-0.428***	-0.151***
	(0.007)	(0.00233)	(0.017)	(0.00592)	(0.022)	(0.00783)
HE Grad.	-1.047***	-0.369***	-0.848***	-0.299***	-0.860***	-0.303***
	(0.008)	(0.00291)	(0.028)	(0.0100)	(0.038)	(0.0135)
HDR	-2.106***	-0.743***	-2.106***	-0.742***	-2.122***	-0.749***
	(0.012)	(0.00424)	(0.012)	(0.00424)	(0.012)	(0.00424)
HDR×Age						
HDR×24-29 years old	-0.827***	-0.292***	-0.829***	-0.292***	-0.823***	-0.290***
	(0.016)	(0.00554)	(0.016)	(0.00554)	(0.016)	(0.00554)
HDR×30-34 years old	-1.748***	-0.616***	-1.750***	-0.617***	-1.743***	-0.615***
	(0.017)	(0.00613)	(0.017)	(0.00613)	(0.017)	(0.00613)
HDR×35-44 years old	-1.879***	-0.663***	-1.881***	-0.663***	-1.870***	-0.660***
	(0.013)	(0.00464)	(0.013)	(0.00464)	(0.013)	(0.00464)
HDR×45-54 years old	-0.941***	-0.332***	-0.942***	-0.332***	-0.933***	-0.329***
	(0.012)	(0.00430)	(0.012)	(0.00430)	(0.012)	(0.00430)
HDR×55 and over	-1.234***	-0.435***	-1.236***	-0.436***	-1.224***	-0.432***
	(0.012)	(0.00424)	(0.012)	(0.00424)	(0.012)	(0.00424)
HDR×Schooling						
HDR×Primary	-0.211***	-0.0742***	-0.210***	-0.0739***	-0.209***	-0.0738***
	(0.014)	(0.00493)	(0.014)	(0.00493)	(0.014)	(0.00493)
HDR×Some Secondary	-0.002	-0.000557	-0.001	-0.000346	0.003	0.000943
	(0.012)	(0.00435)	(0.012)	(0.00435)	(0.012)	(0.00435)
HDR×Secondary Grad.	-0.417***	-0.147***	-0.417***	-0.147***	-0.416***	-0.147***
	(0.011)	(0.00383)	(0.011)	(0.00383)	(0.011)	(0.00383)
HDR×Some HE	-0.341***	-0.120***	-0.342***	-0.121***	-0.340***	-0.120***
	(0.014)	(0.00502)	(0.014)	(0.00502)	(0.014)	(0.00503)
HDR×HE Grad.	-0.446***	-0.157***	-0.446***	-0.157***	-0.441***	-0.156***
	(0.018)	(0.00644)	(0.018)	(0.00644)	(0.018)	(0.00644)
$\bar{Z}_a^{s=o}$	0.134***	0.0472***			0.310***	0.109***
	(0.011)	(0.00396)			(0.015)	(0.00537)
$\bar{Z}_r^{s=o}$	0.368***	0.130***	0.384***	0.135***		
	(0.007)	(0.00256)	(0.007)	(0.00253)		
$\bar{Z}_s^{s=o}$			0.128***	0.0452***	0.131***	0.0464***
			(0.017)	(0.00616)	(0.024)	(0.00845)
Constant	1.683***		1.658***		1.594***	
	(0.009)		(0.016)		(0.016)	
Observations	3,800,372		3,800,372		3,800,372	
Pseudo R^2	0.366		0.366		0.366	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.16: $\mathbb{P}[s = \mathrm{e}] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Age to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.337***	0.0715***	0.741***	0.211***	0.251***	0.0960***
	(0.029)	(0.00809)	(0.008)	(0.00237)	(0.028)	(0.00814)
30 to 34 years old	0.405***	0.0833***	0.932***	0.265***	0.293***	0.115***
	(0.037)	(0.0104)	(0.009)	(0.00255)	(0.037)	(0.0105)
35 to 44 years old	0.296***	0.0493***	0.871***	0.248***	0.173***	0.0843***
	(0.040)	(0.0112)	(0.007)	(0.00213)	(0.039)	(0.0113)
45 to 54 years old	0.092**	-0.00648	0.626***	0.178***	-0.023	0.0262**
	(0.037)	(0.0104)	(0.007)	(0.00210)	(0.037)	(0.0105)
55 years old and over	-0.941***	-0.276***	-0.819***	-0.233***	-0.969***	-0.268***
	(0.010)	(0.00290)	(0.006)	(0.00170)	(0.010)	(0.00291)
Married	1.007***	0.287***	1.007***	0.287***	1.007***	0.287***
	(0.003)	(0.000844)	(0.003)	(0.000844)	(0.003)	(0.000844)
Head of Household	0.249***	0.0711***	0.248***	0.0707***	0.250***	0.0708***
	(0.003)	(0.000863)	(0.003)	(0.000863)	(0.003)	(0.000863)
Schooling						
Completed Primary	-0.050***	0.0443***	-0.060***	-0.0170***	0.156***	-0.0143***
	(0.009)	(0.000983)	(0.009)	(0.00254)	(0.003)	(0.00256)
Some Secondary Ed.	-0.083***	-0.0483***	-0.080***	-0.0226***	-0.170***	-0.0236***
	(0.005)	(0.000865)	(0.005)	(0.00130)	(0.003)	(0.00130)
Secondary School Grad.	0.041***	0.0899***	0.029**	0.00816**	0.316***	0.0118***
	(0.011)	(0.000802)	(0.011)	(0.00322)	(0.003)	(0.00325)
Some HE	-0.225***	-0.0587***	-0.226***	-0.0644***	-0.206***	-0.0640***
	(0.004)	(0.00105)	(0.004)	(0.00107)	(0.004)	(0.00107)
HE Grad.	-0.262***	0.0314***	-0.280***	-0.0796***	0.110***	-0.0746***
	(0.016)	(0.00126)	(0.016)	(0.00443)	(0.004)	(0.00446)
HDR	3.240***	0.922***	3.238***	0.922***	3.239***	0.922***
	(0.008)	(0.00262)	(0.008)	(0.00262)	(0.008)	(0.00262)
HDR×Age						
HDR×24-29 years old	0.478***	0.136***	0.477***	0.136***	0.478***	0.136***
	(0.019)	(0.00527)	(0.019)	(0.00527)	(0.019)	(0.00527)
HDR×30-34 years old	0.842***	0.239***	0.843***	0.240***	0.840***	0.240***
	(0.023)	(0.00639)	(0.023)	(0.00639)	(0.023)	(0.00639)
HDR×35-44 years old	1.525***	0.433***	1.525***	0.434***	1.523***	0.434***
	(0.020)	(0.00546)	(0.020)	(0.00546)	(0.020)	(0.00546)
HDR×45-54 years old	1.004***	0.285***	1.006***	0.286***	1.002***	0.286***
	(0.018)	(0.00499)	(0.018)	(0.00499)	(0.018)	(0.00499)
HDR×55 and over	1.199***	0.341***	1.201***	0.342***	1.197***	0.341***
	(0.012)	(0.00331)	(0.012)	(0.00331)	(0.012)	(0.00331)
$\bar{Z}_a^{s=\mathrm{e}}$	0.331***	0.114***			0.402***	0.0941***
	(0.022)	(0.00635)			(0.022)	(0.00639)
$\bar{Z}_r^{s=\mathrm{e}}$	0.536***	0.0507***	0.560***	0.0447***		
	(0.022)	(0.00382)	(0.021)	(0.00382)		
$\bar{Z}_s^{s=\mathrm{e}}$			0.157***	0.160***	0.178***	0.153***
			(0.013)	(0.00608)	(0.013)	(0.00613)
Constant	-1.890***		-2.083***		-1.798***	
	(0.012)		(0.009)		(0.012)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.487		0.487		0.487	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.17: $\mathbb{P}[s = \text{e}] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Schooling to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.536***	0.134***	0.967***	0.282***	0.459***	0.157***
	(0.027)	(0.00795)	(0.004)	(0.00109)	(0.027)	(0.00800)
30 to 34 years old	0.731***	0.184***	1.294***	0.378***	0.630***	0.213***
	(0.036)	(0.0104)	(0.004)	(0.00124)	(0.036)	(0.0104)
35 to 44 years old	0.803***	0.202***	1.417***	0.413***	0.692***	0.234***
	(0.039)	(0.0113)	(0.004)	(0.00105)	(0.039)	(0.0114)
45 to 54 years old	0.473***	0.108***	1.044***	0.305***	0.370***	0.138***
	(0.036)	(0.0105)	(0.004)	(0.00112)	(0.036)	(0.0106)
55 years old and over	-0.445***	-0.137***	-0.314***	-0.0917***	-0.471***	-0.130***
	(0.009)	(0.00262)	(0.004)	(0.00105)	(0.009)	(0.00264)
Married	0.990***	0.289***	0.990***	0.289***	0.989***	0.289***
	(0.003)	(0.000842)	(0.003)	(0.000842)	(0.003)	(0.000842)
Head of Household	0.241***	0.0706***	0.240***	0.0701***	0.242***	0.0703***
	(0.003)	(0.000874)	(0.003)	(0.000874)	(0.003)	(0.000874)
Schooling						
Completed Primary	0.070***	0.0778***	0.059***	0.0173***	0.267***	0.0203***
	(0.011)	(0.00197)	(0.011)	(0.00309)	(0.007)	(0.00311)
Some Secondary Ed.	0.335***	0.0735***	0.339***	0.0991***	0.252***	0.0978***
	(0.007)	(0.00166)	(0.007)	(0.00194)	(0.006)	(0.00194)
Secondary School Grad.	0.065***	0.0955***	0.051***	0.0150***	0.327***	0.0191***
	(0.012)	(0.00157)	(0.012)	(0.00354)	(0.005)	(0.00356)
Some HE	0.132***	0.0430***	0.131***	0.0381***	0.147***	0.0384***
	(0.007)	(0.00214)	(0.007)	(0.00215)	(0.007)	(0.00215)
HE Grad.	-0.028	0.0961***	-0.047***	-0.0137***	0.329***	-0.00812
	(0.017)	(0.00258)	(0.017)	(0.00504)	(0.009)	(0.00507)
HDR	4.390***	1.279***	4.388***	1.281***	4.385***	1.281***
	(0.009)	(0.00262)	(0.009)	(0.00262)	(0.009)	(0.00262)
HDR×Schooling						
HDR×Primary	-0.312***	-0.0902***	-0.312***	-0.0911***	-0.309***	-0.0909***
	(0.017)	(0.00491)	(0.017)	(0.00491)	(0.017)	(0.00491)
HDR×Some Secondary	-1.131***	-0.330***	-1.131***	-0.330***	-1.131***	-0.330***
	(0.013)	(0.00387)	(0.013)	(0.00387)	(0.013)	(0.00387)
HDR×Secondary Grad.	-0.072***	-0.0198***	-0.073***	-0.0213***	-0.068***	-0.0211***
	(0.013)	(0.00382)	(0.013)	(0.00382)	(0.013)	(0.00382)
HDR×Some HE	-0.946***	-0.274***	-0.948***	-0.277***	-0.940***	-0.276***
	(0.016)	(0.00474)	(0.016)	(0.00475)	(0.016)	(0.00475)
HDR×HE Grad.	-0.612***	-0.178***	-0.613***	-0.179***	-0.609***	-0.178***
	(0.021)	(0.00609)	(0.021)	(0.00610)	(0.021)	(0.00610)
$\bar{Z}_a^{s=\text{e}}$	0.353***	0.122***			0.417***	0.103***
	(0.022)	(0.00646)			(0.022)	(0.00650)
$\bar{Z}_r^{s=\text{e}}$	0.515***	0.0531***	0.542***	0.0484***		
	(0.022)	(0.00390)	(0.021)	(0.00390)		
$\bar{Z}_s^{s=\text{e}}$			0.166***	0.158***	0.182***	0.150***
			(0.013)	(0.00623)	(0.013)	(0.00628)
Constant	-2.338***		-2.544***		-2.255***	
	(0.012)		(0.009)		(0.012)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.486		0.486		0.486	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.18: $\mathbb{P}[s = \mathrm{e}] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with two interaction terms to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.340***	0.0733***	0.746***	0.211***	0.259***	0.0965***
	(0.029)	(0.00807)	(0.008)	(0.00238)	(0.028)	(0.00812)
30 to 34 years old	0.430***	0.0916***	0.960***	0.272***	0.323***	0.122***
	(0.037)	(0.0104)	(0.009)	(0.00256)	(0.037)	(0.0105)
35 to 44 years old	0.332***	0.0611***	0.910***	0.258***	0.215***	0.0942***
	(0.040)	(0.0112)	(0.007)	(0.00215)	(0.039)	(0.0113)
45 to 54 years old	0.131***	0.00630	0.668***	0.189***	0.022	0.0372***
	(0.037)	(0.0104)	(0.007)	(0.00212)	(0.037)	(0.0105)
55 years old and over	-0.884***	-0.258***	-0.761***	-0.216***	-0.910***	-0.251***
	(0.011)	(0.00296)	(0.006)	(0.00182)	(0.010)	(0.00297)
Married	1.009***	0.286***	1.008***	0.286***	1.008***	0.286***
	(0.003)	(0.000842)	(0.003)	(0.000842)	(0.003)	(0.000842)
Head of Household	0.249***	0.0709***	0.248***	0.0704***	0.250***	0.0706***
	(0.003)	(0.000859)	(0.003)	(0.000859)	(0.003)	(0.000859)
Schooling						
Completed Primary	-0.037***	0.0463***	-0.047***	-0.0132***	0.163***	-0.0106***
	(0.011)	(0.00198)	(0.011)	(0.00305)	(0.007)	(0.00307)
Some Secondary Ed.	0.158***	0.0208***	0.161***	0.0458***	0.073***	0.0448***
	(0.007)	(0.00175)	(0.007)	(0.00200)	(0.006)	(0.00200)
Secondary School Grad.	-0.067***	0.0568***	-0.079***	-0.0225***	0.200***	-0.0189***
	(0.012)	(0.00162)	(0.012)	(0.00349)	(0.006)	(0.00351)
Some HE	-0.041***	-0.00696***	-0.042***	-0.0118***	-0.025***	-0.0115***
	(0.008)	(0.00219)	(0.008)	(0.00220)	(0.008)	(0.00220)
HE Grad.	-0.088***	0.0784***	-0.105***	-0.0298***	0.276***	-0.0249***
	(0.018)	(0.00256)	(0.017)	(0.00494)	(0.009)	(0.00497)
HDR	3.445***	0.976***	3.443***	0.976***	3.441***	0.977***
	(0.014)	(0.00416)	(0.014)	(0.00416)	(0.014)	(0.00416)
HDR×Age						
HDR×24-29 years old	0.482***	0.137***	0.482***	0.137***	0.482***	0.137***
	(0.019)	(0.00536)	(0.019)	(0.00536)	(0.019)	(0.00536)
HDR×30-34 years old	0.789***	0.223***	0.789***	0.224***	0.786***	0.224***
	(0.023)	(0.00639)	(0.023)	(0.00639)	(0.023)	(0.00639)
HDR×35-44 years old	1.428***	0.404***	1.428***	0.405***	1.426***	0.405***
	(0.020)	(0.00552)	(0.020)	(0.00552)	(0.020)	(0.00552)
HDR×45-54 years old	0.902***	0.255***	0.904***	0.256***	0.900***	0.256***
	(0.018)	(0.00506)	(0.018)	(0.00506)	(0.018)	(0.00506)
HDR×55 and over	1.059***	0.300***	1.061***	0.301***	1.058***	0.300***
	(0.013)	(0.00378)	(0.013)	(0.00378)	(0.013)	(0.00378)
HDR×Schooling						
HDR×Primary	-0.020	-0.00478	-0.020	-0.00559	-0.017	-0.00557
	(0.018)	(0.00499)	(0.018)	(0.00499)	(0.018)	(0.00499)
HDR×Some Secondary	-0.646***	-0.183***	-0.646***	-0.183***	-0.647***	-0.183***
	(0.015)	(0.00423)	(0.015)	(0.00423)	(0.015)	(0.00423)
HDR×Secondary Grad.	0.308***	0.0886***	0.308***	0.0874***	0.312***	0.0874***
	(0.014)	(0.00402)	(0.014)	(0.00402)	(0.014)	(0.00402)
HDR×Some HE	-0.472***	-0.132***	-0.473***	-0.134***	-0.466***	-0.134***
	(0.018)	(0.00503)	(0.018)	(0.00503)	(0.018)	(0.00503)
HDR×HE Grad.	-0.460***	-0.130***	-0.460***	-0.131***	-0.458***	-0.130***
	(0.022)	(0.00611)	(0.022)	(0.00611)	(0.022)	(0.00611)
$\bar{Z}_a^{s=\mathrm{e}}$	0.332***	0.113***			0.400***	0.0942***
	(0.022)	(0.00633)			(0.022)	(0.00637)
$\bar{Z}_r^{s=\mathrm{e}}$	0.525***	0.0510***	0.549***	0.0455***		
	(0.022)	(0.00380)	(0.021)	(0.00380)		
$\bar{Z}_s^{s=\mathrm{e}}$			0.160***	0.156***	0.180***	0.149***
			(0.013)	(0.00607)	(0.013)	(0.00612)
Constant	-1.967***		-2.162***		-1.879***	
	(0.013)		(0.010)		(0.013)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.488		0.488		0.488	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.19: $\mathbb{P}[s=u] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Age to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.753***	0.0445***	0.803***	0.0457***	0.782***	0.0431***
	(0.009)	(0.000523)	(0.008)	(0.000493)	(0.009)	(0.000549)
30 to 34 years old	0.813***	0.0458***	0.800***	0.0455***	0.805***	0.0466***
	(0.009)	(0.000524)	(0.009)	(0.000522)	(0.009)	(0.000529)
35 to 44 years old	0.838***	0.0450***	0.759***	0.0432***	0.791***	0.0480***
	(0.010)	(0.000506)	(0.008)	(0.000426)	(0.009)	(0.000565)
45 to 54 years old	0.682***	0.0350***	0.570***	0.0324***	0.615***	0.0391***
	(0.012)	(0.000584)	(0.008)	(0.000439)	(0.010)	(0.000681)
55 years old and over	-0.195***	-0.0195***	-0.443***	-0.0252***	-0.343***	-0.0112***
	(0.021)	(0.000944)	(0.007)	(0.000415)	(0.017)	(0.00122)
Married	-0.186***	-0.0107***	-0.189***	-0.0107***	-0.189***	-0.0107***
	(0.004)	(0.000211)	(0.004)	(0.000211)	(0.004)	(0.000213)
Head of Household	-0.322***	-0.0180***	-0.317***	-0.0180***	-0.317***	-0.0184***
	(0.004)	(0.000218)	(0.004)	(0.000218)	(0.004)	(0.000219)
Schooling						
Completed Primary	0.013	0.00695***	0.072***	0.00407***	0.122***	0.000720
	(0.008)	(0.000276)	(0.007)	(0.000392)	(0.005)	(0.000476)
Some Secondary Ed.	-0.101***	-0.000264	-0.052***	-0.00294***	-0.005	-0.00579***
	(0.008)	(0.000253)	(0.006)	(0.000361)	(0.004)	(0.000442)
Secondary School Grad.	0.137***	0.0181***	0.232***	0.0132***	0.319***	0.00785***
	(0.012)	(0.000224)	(0.009)	(0.000525)	(0.004)	(0.000700)
Some HE	-0.037***	0.00430***	0.020***	0.00111***	0.076***	-0.00210***
	(0.009)	(0.000295)	(0.007)	(0.000424)	(0.005)	(0.000523)
HE Grad.	0.164***	0.0154***	0.217***	0.0123***	0.270***	0.00941***
	(0.009)	(0.000332)	(0.008)	(0.000441)	(0.006)	(0.000531)
HDR	-1.376***	-0.0781***	-1.372***	-0.0780***	-1.372***	-0.0789***
	(0.011)	(0.000651)	(0.011)	(0.000651)	(0.011)	(0.000654)
HDR×Age						
HDR×24-29 years old	-1.415***	-0.0806***	-1.417***	-0.0806***	-1.417***	-0.0811***
	(0.021)	(0.00122)	(0.021)	(0.00122)	(0.021)	(0.00123)
HDR×30-34 years old	-1.928***	-0.110***	-1.928***	-0.110***	-1.928***	-0.110***
	(0.025)	(0.00145)	(0.026)	(0.00145)	(0.026)	(0.00146)
HDR×35-44 years old	-2.589***	-0.147***	-2.590***	-0.147***	-2.589***	-0.148***
	(0.023)	(0.00126)	(0.023)	(0.00126)	(0.023)	(0.00127)
HDR×45-54 years old	-1.627***	-0.0926***	-1.628***	-0.0926***	-1.628***	-0.0932***
	(0.022)	(0.00120)	(0.022)	(0.00120)	(0.022)	(0.00121)
HDR×55 and over	0.443***	0.0251***	0.441***	0.0251***	0.442***	0.0254***
	(0.017)	(0.000991)	(0.017)	(0.000990)	(0.017)	(0.000996)
$\bar{Z}_a^{s=u}$	0.367***	0.00868***			0.153***	0.0210***
	(0.030)	(0.00128)			(0.022)	(0.00174)
$\bar{Z}_r^{s=u}$	0.465***	0.0427***	0.212***	0.0414***		
	(0.028)	(0.000856)	(0.021)	(0.000840)		
$\bar{Z}_s^{s=u}$			0.727***	0.0121***	0.750***	0.0267***
			(0.015)	(0.00117)	(0.015)	(0.00162)
Constant	0.318***		0.509***		0.368***	
	(0.035)		(0.035)		(0.030)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.172		0.174		0.174	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.20: $\mathbb{P}[s=u] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Schooling to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.285***	0.0191***	0.339***	0.0204***	0.316***	0.0173***
	(0.006)	(0.000322)	(0.004)	(0.000264)	(0.005)	(0.000364)
30 to 34 years old	0.211***	0.0122***	0.197***	0.0119***	0.202***	0.0128***
	(0.005)	(0.000293)	(0.005)	(0.000290)	(0.005)	(0.000298)
35 to 44 years old	0.112***	0.00374***	0.027***	0.00165***	0.062***	0.00677***
	(0.008)	(0.000387)	(0.004)	(0.000262)	(0.006)	(0.000467)
45 to 54 years old	0.185***	0.00685***	0.065***	0.00392***	0.114***	0.0112***
	(0.010)	(0.000496)	(0.005)	(0.000291)	(0.008)	(0.000615)
55 years old and over	-0.222***	-0.0229***	-0.487***	-0.0293***	-0.379***	-0.0135***
	(0.020)	(0.000944)	(0.005)	(0.000317)	(0.016)	(0.00124)
Married	-0.172***	-0.0105***	-0.174***	-0.0105***	-0.174***	-0.0105***
	(0.004)	(0.000219)	(0.004)	(0.000219)	(0.004)	(0.000220)
Head of Household	-0.306***	-0.0182***	-0.302***	-0.0182***	-0.302***	-0.0186***
	(0.004)	(0.000227)	(0.004)	(0.000227)	(0.004)	(0.000228)
Schooling						
Completed Primary	0.216***	0.0194***	0.270***	0.0163***	0.321***	0.0131***
	(0.010)	(0.000479)	(0.009)	(0.000562)	(0.008)	(0.000630)
Some Secondary Ed.	-0.073***	0.00119***	-0.027***	-0.00165***	0.020***	-0.00444***
	(0.009)	(0.000420)	(0.008)	(0.000499)	(0.007)	(0.000566)
Secondary School Grad.	0.259***	0.0262***	0.348***	0.0210***	0.435***	0.0157***
	(0.013)	(0.000383)	(0.010)	(0.000627)	(0.006)	(0.000792)
Some HE	-0.135***	-0.00163***	-0.083***	-0.00500***	-0.027***	-0.00822***
	(0.011)	(0.000503)	(0.010)	(0.000595)	(0.008)	(0.000675)
HE Grad.	0.263***	0.0220***	0.311***	0.0188***	0.365***	0.0160***
	(0.012)	(0.000599)	(0.011)	(0.000671)	(0.010)	(0.000742)
HDR	-1.925***	-0.116***	-1.920***	-0.116***	-1.920***	-0.117***
	(0.014)	(0.000782)	(0.014)	(0.000782)	(0.014)	(0.000786)
HDR×Schooling						
HDR×Primary	-0.805***	-0.0485***	-0.803***	-0.0484***	-0.804***	-0.0489***
	(0.025)	(0.00151)	(0.025)	(0.00151)	(0.025)	(0.00152)
HDR×Some Secondary	-0.190***	-0.0116***	-0.192***	-0.0116***	-0.192***	-0.0116***
	(0.021)	(0.00127)	(0.021)	(0.00127)	(0.021)	(0.00128)
HDR×Secondary Grad.	-0.472***	-0.0286***	-0.474***	-0.0286***	-0.474***	-0.0287***
	(0.018)	(0.00110)	(0.018)	(0.00110)	(0.018)	(0.00110)
HDR×Some HE	0.212***	0.0127***	0.210***	0.0126***	0.210***	0.0129***
	(0.023)	(0.00138)	(0.023)	(0.00138)	(0.023)	(0.00138)
HDR×HE Grad.	-0.456***	-0.0277***	-0.458***	-0.0276***	-0.459***	-0.0277***
	(0.027)	(0.00163)	(0.027)	(0.00163)	(0.027)	(0.00164)
$\bar{Z}_a^{s=u}$	0.391***	0.00986***			0.164***	0.0238***
	(0.030)	(0.00134)			(0.022)	(0.00182)
$\bar{Z}_r^{s=u}$	0.448***	0.0451***	0.212***	0.0440***		
	(0.028)	(0.000897)	(0.020)	(0.000880)		
$\bar{Z}_s^{s=u}$			0.730***	0.0128***	0.748***	0.0272***
			(0.015)	(0.00123)	(0.015)	(0.00170)
Constant	0.526***		0.717***		0.586***	
	(0.035)		(0.035)		(0.030)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.155		0.156		0.156	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.21: $\mathbb{P}[s=u] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with two interaction terms to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	0.780***	0.0458***	0.831***	0.0470***	0.810***	0.0445***
	(0.009)	(0.000529)	(0.009)	(0.000500)	(0.009)	(0.000555)
30 to 34 years old	0.820***	0.0459***	0.807***	0.0456***	0.812***	0.0467***
	(0.009)	(0.000526)	(0.009)	(0.000524)	(0.009)	(0.000530)
35 to 44 years old	0.841***	0.0449***	0.761***	0.0431***	0.793***	0.0479***
	(0.010)	(0.000508)	(0.008)	(0.000430)	(0.009)	(0.000566)
45 to 54 years old	0.683***	0.0348***	0.570***	0.0322***	0.615***	0.0389***
	(0.012)	(0.000587)	(0.008)	(0.000444)	(0.010)	(0.000682)
55 years old and over	-0.183***	-0.0189***	-0.433***	-0.0245***	-0.334***	-0.0104***
	(0.021)	(0.000949)	(0.007)	(0.000432)	(0.017)	(0.00122)
Married	-0.189***	-0.0108***	-0.192***	-0.0108***	-0.192***	-0.0108***
	(0.004)	(0.000211)	(0.004)	(0.000211)	(0.004)	(0.000212)
Head of Household	-0.322***	-0.0180***	-0.318***	-0.0180***	-0.318***	-0.0184***
	(0.004)	(0.000217)	(0.004)	(0.000217)	(0.004)	(0.000219)
Schooling						
Completed Primary	0.133***	0.0136***	0.191***	0.0108***	0.241***	0.00760***
	(0.011)	(0.000462)	(0.010)	(0.000539)	(0.008)	(0.000604)
Some Secondary Ed.	-0.067***	0.00159***	-0.018**	-0.00103**	0.028***	-0.00381***
	(0.010)	(0.000419)	(0.009)	(0.000490)	(0.007)	(0.000553)
Secondary School Grad.	0.204***	0.0217***	0.298***	0.0169***	0.383***	0.0116***
	(0.013)	(0.000368)	(0.011)	(0.000599)	(0.006)	(0.000755)
Some HE	-0.180***	-0.00392***	-0.124***	-0.00703***	-0.069***	-0.0103***
	(0.012)	(0.000498)	(0.010)	(0.000582)	(0.009)	(0.000658)
HE Grad.	0.121***	0.0128***	0.173***	0.00979***	0.226***	0.00692***
	(0.012)	(0.000576)	(0.011)	(0.000644)	(0.010)	(0.000711)
HDR	-1.273***	-0.0718***	-1.268***	-0.0717***	-1.269***	-0.0725***
	(0.020)	(0.00112)	(0.020)	(0.00112)	(0.020)	(0.00113)
HDR×Age						
HDR×24-29 years old	-1.506***	-0.0853***	-1.507***	-0.0853***	-1.507***	-0.0858***
	(0.022)	(0.00125)	(0.022)	(0.00125)	(0.022)	(0.00126)
HDR×30-34 years old	-1.956***	-0.111***	-1.957***	-0.111***	-1.956***	-0.111***
	(0.026)	(0.00146)	(0.026)	(0.00146)	(0.026)	(0.00147)
HDR×35-44 years old	-2.608***	-0.148***	-2.610***	-0.148***	-2.610***	-0.149***
	(0.024)	(0.00128)	(0.024)	(0.00128)	(0.024)	(0.00129)
HDR×45-54 years old	-1.627***	-0.0921***	-1.628***	-0.0921***	-1.628***	-0.0927***
	(0.022)	(0.00122)	(0.022)	(0.00122)	(0.022)	(0.00123)
HDR×55 and over	0.399***	0.0225***	0.398***	0.0225***	0.398***	0.0228***
	(0.019)	(0.00109)	(0.019)	(0.00109)	(0.019)	(0.00110)
HDR×Schooling						
HDR×Primary	-0.456***	-0.0257***	-0.453***	-0.0256***	-0.454***	-0.0260***
	(0.026)	(0.00145)	(0.026)	(0.00145)	(0.026)	(0.00146)
HDR×Some Secondary	-0.137***	-0.00768***	-0.136***	-0.00770***	-0.136***	-0.00779***
	(0.023)	(0.00128)	(0.023)	(0.00128)	(0.023)	(0.00129)
HDR×Secondary Grad.	-0.228***	-0.0129***	-0.229***	-0.0129***	-0.228***	-0.0130***
	(0.019)	(0.00109)	(0.019)	(0.00109)	(0.019)	(0.00110)
HDR×Some HE	0.445***	0.0252***	0.444***	0.0251***	0.445***	0.0254***
	(0.024)	(0.00139)	(0.024)	(0.00139)	(0.024)	(0.00139)
HDR×HE Grad.	0.115***	0.00644***	0.115***	0.00650***	0.114***	0.00655***
	(0.027)	(0.00155)	(0.027)	(0.00155)	(0.027)	(0.00156)
$\bar{Z}_a^{s=u}$	0.370***	0.00853***			0.151***	0.0211***
	(0.030)	(0.00127)			(0.023)	(0.00173)
$\bar{Z}_r^{s=u}$	0.459***	0.0424***	0.209***	0.0411***		
	(0.028)	(0.000853)	(0.021)	(0.000837)		
$\bar{Z}_s^{s=u}$			0.727***	0.0118***	0.749***	0.0262***
			(0.015)	(0.00117)	(0.015)	(0.00161)
Constant	0.287***		0.480***		0.341***	
	(0.035)		(0.036)		(0.031)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.173		0.175		0.175	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.22: $\mathbb{P}[s=o] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Age to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.439***	-0.120***	-1.162***	-0.257***	-0.543***	-0.0970***
	(0.039)	(0.00864)	(0.008)	(0.00185)	(0.039)	(0.00864)
30 to 34 years old	-0.485***	-0.137***	-1.440***	-0.318***	-0.622***	-0.107***
	(0.051)	(0.0114)	(0.010)	(0.00224)	(0.051)	(0.0114)
35 to 44 years old	-0.260***	-0.0884***	-1.239***	-0.274***	-0.400***	-0.0575***
	(0.052)	(0.0116)	(0.007)	(0.00177)	(0.052)	(0.0116)
45 to 54 years old	0.210***	0.0201**	-0.631***	-0.139***	0.091**	0.0465***
	(0.045)	(0.00996)	(0.007)	(0.00162)	(0.045)	(0.00995)
55 years old and over	1.292***	0.283***	1.182***	0.261***	1.278***	0.286***
	(0.008)	(0.00180)	(0.006)	(0.00126)	(0.008)	(0.00180)
Married	-1.000***	-0.221***	-1.000***	-0.221***	-0.999***	-0.221***
	(0.003)	(0.000717)	(0.003)	(0.000717)	(0.003)	(0.000717)
Head of Household	-0.197***	-0.0439***	-0.196***	-0.0433***	-0.199***	-0.0435***
	(0.003)	(0.000735)	(0.003)	(0.000735)	(0.003)	(0.000735)
Schooling						
Completed Primary	0.127***	-0.0451***	0.104***	0.0230***	-0.204***	0.0280***
	(0.011)	(0.000801)	(0.011)	(0.00243)	(0.004)	(0.00243)
Some Secondary Ed.	0.094***	0.0386***	0.100***	0.0220***	0.175***	0.0207***
	(0.004)	(0.000686)	(0.004)	(0.000879)	(0.003)	(0.000875)
Secondary School Grad.	-0.038**	-0.117***	-0.071***	-0.0156***	-0.530***	-0.00841**
	(0.016)	(0.000668)	(0.016)	(0.00348)	(0.003)	(0.00348)
Some HE	0.259***	0.0418***	0.255***	0.0565***	0.189***	0.0573***
	(0.004)	(0.000829)	(0.004)	(0.000970)	(0.004)	(0.000970)
HE Grad.	0.348***	-0.0546***	0.309***	0.0683***	-0.247***	0.0768***
	(0.019)	(0.00107)	(0.019)	(0.00428)	(0.005)	(0.00428)
HDR	-2.350***	-0.519***	-2.347***	-0.519***	-2.347***	-0.519***
	(0.008)	(0.00186)	(0.008)	(0.00186)	(0.008)	(0.00186)
HDR×Age						
HDR×24-29 years old	-0.020	-0.00408	-0.017	-0.00368	-0.018	-0.00439
	(0.018)	(0.00405)	(0.018)	(0.00405)	(0.018)	(0.00405)
HDR×30-34 years old	-0.285***	-0.0626***	-0.283***	-0.0625***	-0.283***	-0.0630***
	(0.025)	(0.00553)	(0.025)	(0.00553)	(0.025)	(0.00553)
HDR×35-44 years old	-0.912***	-0.201***	-0.910***	-0.201***	-0.910***	-0.202***
	(0.021)	(0.00454)	(0.021)	(0.00454)	(0.021)	(0.00454)
HDR×45-54 years old	-1.125***	-0.248***	-1.127***	-0.249***	-1.124***	-0.249***
	(0.018)	(0.00389)	(0.018)	(0.00389)	(0.018)	(0.00389)
HDR×55 and over	-1.793***	-0.396***	-1.796***	-0.397***	-1.792***	-0.396***
	(0.011)	(0.00241)	(0.011)	(0.00241)	(0.011)	(0.00241)
$\bar{Z}_a^{s=o}$	0.503***	0.0953***			0.431***	0.111***
	(0.027)	(0.00589)			(0.027)	(0.00589)
$\bar{Z}_r^{s=o}$	0.703***	0.0432***	0.655***	0.0379***		
	(0.022)	(0.00323)	(0.022)	(0.00324)		
$\bar{Z}_s^{s=o}$			0.171***	0.145***	0.195***	0.155***
			(0.015)	(0.00488)	(0.015)	(0.00487)
Constant	1.475***		1.643***		1.361***	
	(0.010)		(0.012)		(0.011)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.490		0.490		0.490	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.23: $\mathbb{P}[s=o] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with interaction HDR×Schooling to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.373***	-0.106***	-1.200***	-0.264***	-0.481***	-0.0822***
	(0.038)	(0.00837)	(0.004)	(0.000939)	(0.038)	(0.00836)
30 to 34 years old	-0.511***	-0.144***	-1.604***	-0.353***	-0.652***	-0.113***
	(0.050)	(0.0111)	(0.005)	(0.00116)	(0.050)	(0.0111)
35 to 44 years old	-0.474***	-0.136***	-1.595***	-0.351***	-0.618***	-0.104***
	(0.051)	(0.0113)	(0.004)	(0.000913)	(0.051)	(0.0113)
45 to 54 years old	-0.091**	-0.0471***	-1.055***	-0.232***	-0.214***	-0.0200**
	(0.044)	(0.00974)	(0.004)	(0.000933)	(0.044)	(0.00974)
55 years old and over	0.637***	0.137***	0.510***	0.112***	0.623***	0.140***
	(0.007)	(0.00152)	(0.004)	(0.000822)	(0.007)	(0.00152)
Married	-0.978***	-0.215***	-0.978***	-0.215***	-0.977***	-0.215***
	(0.003)	(0.000679)	(0.003)	(0.000679)	(0.003)	(0.000679)
Head of Household	-0.208***	-0.0462***	-0.207***	-0.0456***	-0.210***	-0.0459***
	(0.003)	(0.000721)	(0.003)	(0.000721)	(0.003)	(0.000721)
Schooling						
Completed Primary	-0.143***	-0.104***	-0.169***	-0.0372***	-0.473***	-0.0315***
	(0.012)	(0.00145)	(0.012)	(0.00269)	(0.007)	(0.00269)
Some Secondary Ed.	-0.422***	-0.0753***	-0.416***	-0.0917***	-0.342***	-0.0930***
	(0.006)	(0.00119)	(0.006)	(0.00131)	(0.005)	(0.00131)
Secondary School Grad.	-0.224***	-0.157***	-0.261***	-0.0576***	-0.714***	-0.0493***
	(0.016)	(0.00117)	(0.016)	(0.00357)	(0.005)	(0.00357)
Some HE	-0.148***	-0.0476***	-0.154***	-0.0338***	-0.216***	-0.0327***
	(0.007)	(0.00152)	(0.007)	(0.00159)	(0.007)	(0.00160)
HE Grad.	0.057***	-0.118***	0.012	0.00260	-0.538***	0.0124***
	(0.021)	(0.00196)	(0.021)	(0.00454)	(0.009)	(0.00454)
HDR	-3.921***	-0.862***	-3.920***	-0.863***	-3.916***	-0.864***
	(0.008)	(0.00200)	(0.008)	(0.00200)	(0.008)	(0.00200)
HDR×Schooling						
HDR×Primary	0.788***	0.173***	0.789***	0.174***	0.787***	0.174***
	(0.016)	(0.00357)	(0.016)	(0.00357)	(0.016)	(0.00357)
HDR×Some Secondary	1.455***	0.321***	1.456***	0.321***	1.456***	0.320***
	(0.012)	(0.00277)	(0.012)	(0.00277)	(0.012)	(0.00277)
HDR×Secondary Grad.	0.598***	0.131***	0.599***	0.132***	0.596***	0.132***
	(0.013)	(0.00283)	(0.013)	(0.00283)	(0.013)	(0.00283)
HDR×Some HE	1.172***	0.257***	1.173***	0.258***	1.165***	0.258***
	(0.015)	(0.00339)	(0.015)	(0.00340)	(0.015)	(0.00340)
HDR×HE Grad.	0.844***	0.186***	0.845***	0.186***	0.845***	0.186***
	(0.021)	(0.00458)	(0.021)	(0.00458)	(0.021)	(0.00458)
$\bar{Z}_a^{s=o}$	0.576***	0.110***			0.502***	0.127***
	(0.026)	(0.00580)			(0.026)	(0.00580)
$\bar{Z}_r^{s=o}$	0.702***	0.0430***	0.648***	0.0384***		
	(0.022)	(0.00320)	(0.022)	(0.00321)		
$\bar{Z}_s^{s=o}$			0.174***	0.143***	0.195***	0.155***
			(0.015)	(0.00484)	(0.015)	(0.00483)
Constant	2.039***		2.221***		1.925***	
	(0.010)		(0.012)		(0.011)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.486		0.486		0.486	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.24: $\mathbb{P}[s=o] = \bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$ with two interaction terms to men

Covariates	(1)		(2)		(3)	
	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$	β	$\partial\bar{y}/\partial x$
Age						
24 to 29 years old	-0.442***	-0.120***	-1.166***	-0.257***	-0.544***	-0.0974***
	(0.039)	(0.00862)	(0.008)	(0.00187)	(0.039)	(0.00862)
30 to 34 years old	-0.501***	-0.140***	-1.459***	-0.321***	-0.635***	-0.110***
	(0.051)	(0.0114)	(0.010)	(0.00223)	(0.051)	(0.0113)
35 to 44 years old	-0.290***	-0.0939***	-1.271***	-0.280***	-0.426***	-0.0638***
	(0.052)	(0.0115)	(0.007)	(0.00179)	(0.052)	(0.0115)
45 to 54 years old	0.174***	0.0128	-0.668***	-0.147***	0.058	0.0384***
	(0.045)	(0.00993)	(0.007)	(0.00164)	(0.045)	(0.00992)
55 years old and over	1.230***	0.268***	1.120***	0.247***	1.217***	0.271***
	(0.008)	(0.00185)	(0.006)	(0.00133)	(0.008)	(0.00185)
Married	-0.999***	-0.220***	-0.999***	-0.220***	-0.998***	-0.220***
	(0.003)	(0.000717)	(0.003)	(0.000717)	(0.003)	(0.000717)
Head of Household	-0.197***	-0.0438***	-0.196***	-0.0433***	-0.199***	-0.0435***
	(0.003)	(0.000732)	(0.003)	(0.000732)	(0.003)	(0.000732)
Schooling						
Completed Primary	0.057***	-0.0594***	0.035***	0.00781***	-0.270***	0.0126***
	(0.012)	(0.00151)	(0.012)	(0.00274)	(0.007)	(0.00274)
Some Secondary Ed.	-0.129***	-0.0109***	-0.124***	-0.0272***	-0.050***	-0.0285***
	(0.006)	(0.00131)	(0.006)	(0.00142)	(0.006)	(0.00141)
Secondary School Grad.	-0.001	-0.107***	-0.033**	-0.00733**	-0.488***	-0.000231
	(0.016)	(0.00124)	(0.016)	(0.00362)	(0.006)	(0.00361)
Some HE	0.161***	0.0208***	0.158***	0.0348***	0.094***	0.0356***
	(0.008)	(0.00162)	(0.008)	(0.00169)	(0.007)	(0.00169)
HE Grad.	0.189***	-0.0883***	0.150***	0.0331***	-0.401***	0.0416***
	(0.021)	(0.00202)	(0.021)	(0.00459)	(0.009)	(0.00459)
HDR	-2.595***	-0.571***	-2.592***	-0.571***	-2.592***	-0.572***
	(0.013)	(0.00306)	(0.013)	(0.00306)	(0.013)	(0.00306)
HDR×Age						
HDR×24-29 years old	-0.020	-0.00395	-0.017	-0.00373	-0.018	-0.00445
	(0.019)	(0.00414)	(0.019)	(0.00414)	(0.019)	(0.00414)
HDR×30-34 years old	-0.250***	-0.0546***	-0.247***	-0.0544***	-0.248***	-0.0550***
	(0.025)	(0.00552)	(0.025)	(0.00552)	(0.025)	(0.00552)
HDR×35-44 years old	-0.830***	-0.183***	-0.828***	-0.182***	-0.829***	-0.183***
	(0.021)	(0.00457)	(0.021)	(0.00457)	(0.021)	(0.00457)
HDR×45-54 years old	-1.031***	-0.227***	-1.034***	-0.228***	-1.031***	-0.227***
	(0.018)	(0.00394)	(0.018)	(0.00394)	(0.018)	(0.00394)
HDR×55 and over	-1.637***	-0.360***	-1.640***	-0.361***	-1.636***	-0.361***
	(0.012)	(0.00274)	(0.012)	(0.00275)	(0.012)	(0.00275)
HDR×Schooling						
HDR×Primary	0.190***	0.0414***	0.190***	0.0418***	0.188***	0.0420***
	(0.017)	(0.00376)	(0.017)	(0.00377)	(0.017)	(0.00377)
HDR×Some Secondary	0.613***	0.135***	0.612***	0.135***	0.614***	0.135***
	(0.014)	(0.00315)	(0.014)	(0.00315)	(0.014)	(0.00315)
HDR×Secondary Grad.	-0.111***	-0.0249***	-0.111***	-0.0245***	-0.113***	-0.0245***
	(0.014)	(0.00305)	(0.014)	(0.00306)	(0.014)	(0.00306)
HDR×Some HE	0.269***	0.0579***	0.269***	0.0592***	0.263***	0.0593***
	(0.017)	(0.00375)	(0.017)	(0.00375)	(0.017)	(0.00375)
HDR×HE Grad.	0.435***	0.0960***	0.436***	0.0960***	0.436***	0.0959***
	(0.022)	(0.00478)	(0.022)	(0.00478)	(0.022)	(0.00478)
$\bar{Z}_a^{s=o}$	0.504***	0.0956***			0.434***	0.111***
	(0.027)	(0.00587)			(0.027)	(0.00587)
$\bar{Z}_r^{s=o}$	0.697***	0.0430***	0.650***	0.0379***		
	(0.022)	(0.00322)	(0.022)	(0.00323)		
$\bar{Z}_s^{s=o}$			0.172***	0.143***	0.195***	0.153***
			(0.015)	(0.00487)	(0.015)	(0.00486)
Constant	1.563***		1.733***		1.451***	
	(0.011)		(0.013)		(0.012)	
Observations	3,482,735		3,482,735		3,482,735	
Pseudo R^2	0.491		0.491		0.491	

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Base outcome: Age: 15 to 24 years old. Schooling: Some Primary Education.

Table B.25: Adjusted Prediction for $\bar{X}\beta + \lambda\bar{Z}^s + \mu\bar{Z}^s$

Covariates	**Women**			**Men**		
	$s = \mathrm{e}$	$s = u$	$s = o$	$s = \mathrm{e}$	$s = u$	$s = o$
Prediction for $\mathbb{P}(\mathrm{e}, \mathrm{age})$	5.496***			1.415***		
	(0.519)			(0.277)		
Prediction for $\mathbb{P}(\mathrm{e}, \mathrm{region})$	-4.866***			-1.139***		
	(0.512)			(0.285)		
Prediction for $\mathbb{P}(u, \mathrm{age})$		0.108			1,130.421	
		(0.202)			(2,318.782)	
Prediction for $\mathbb{P}(u, \mathrm{region})$		-0.080			-373.102	
		(0.184)			(4,523.265)	
Prediction for $\mathbb{P}(o, \mathrm{age})$			5.915***			-0.388
			(0.363)			(0.259)
Prediction for $\mathbb{P}(o, \mathrm{region})$			-5.153***			1.021**
			(0.342)			(0.492)
Constant	-0.034	0.023**	0.329***	0.612***	0.048***	0.302***
	(0.022)	(0.010)	(0.009)	(0.052)	(0.001)	(0.023)
Observations	84	84	84	84	84	84
R^2	0.935	0.0200	0.947	0.587	0.283	0.280

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1