Informality and Mobility: Evidence from Russian Panel Data*

Fabián Slonimczyk and Vladimir Gimpelson†

September 23, 2013

Abstract

Informality is a defining characteristic of labor markets in developing and transition countries. This paper analyzes patterns of mobility across different forms of formal and informal employment in Russia. Using the RLMS household panel we estimate a dynamic multinomial logit model with individual heterogeneity and correct for the initial conditions problem. Simulations show that structural state dependence is weak and that transition rates from informal to formal employment are not lower than from other origin states. These results lend support to the integrated view of the labor market.

JEL classification: J6

Key words: informality ; labor mobility; initial conditions problem; state dependence

*This paper has been presented at the following venues: Higher School of Economics, First RLMS International Users' Conference, Italian Labor Economics Association. We thank participants for helpful comments. We are grateful to Anna Yurko for useful suggestions to improve the manuscript. Fabian Slonimczyk acknowledges support from NRU Higher School of Economics Academic Fund Program 2012–2013 (research grant No.11-01-0165). Vladimir Gimpelson acknowledges support from NRU Higher School of Economics Basic Research Program.

†National Research University Higher School of Economics, Moscow; fslonimczyk@hse.ru and vladim@hse.ru.

1 Introduction

Widespread informality is a salient feature of transition and emerging economies. It characterizes a broad set of economic activities, including the operations of the myriad small-scale enterprizes that are fully run by families or single individuals, but also informally hired employees working for otherwise formal firms.[1] Overall, in these countries a large share of the population is informally employed.[2]

There are multiple reasons why sprawling informality is seen as a negative phenomenon. First, the informal economy operates largely beyond State regulation and tax evasion is the norm rather than the exception. Second, the need to keep activities undetected leads to multiple inefficiencies including suboptimal scale, low investment rates, use of primitive technologies, and close to zero productive innovation. Third, earnings in some informal activities are low and irregular, linking informality and poverty. Finally, informal workers face higher risks of social exclusion because they are not covered by any kind of safety net (except maybe the one provided by family and friends).

Various dimensions of this phenomenon have been thoroughly examined, especially for Latin American countries (Perry et al., 2007). However, the drivers and implications of informality in transition countries remain largely unexplored (for exceptions, see Lehmann et al., 2012, Slonimczyk, 2012, Gimpelson and Kapeliushnikov, 2013b)

An important issue is whether the same individuals participate in the informal sector year after year, or if instead the incidence of informality is spread more equally across the labor force. If informality is a persistent state, most of the negative effects associated with it would be suffered by a limited set of individuals. If the opposite, everyone in the population faces roughly equal chances of experiencing a spell of informal employment.

A closely related question is whether transitions out of informality lead to formal jobs or if, on the contrary, informal workers end up dropping out of the labor force. Does informal employment function as a "stepping stone" toward formal positions or is it a "dead end" without exit to better jobs? A priori there are good arguments for both views. There are several reasons to think that the probability of finding a formal job might be positively related to informal work experience. First, informal jobs might contribute to general human capital, increasing the worker's value in the market. Second, workers might gain in terms of an expanded social and professional network (compared to a non-employment alternative). This could result in better information on existing job vacancies and a relatively higher rate of arrival for offers from the formal sector. Third, some firms might use informal positions as a screening device and later offer regular positions to the best informal trainees. Finally, informal work could signal higher levels of ability or other unobservable traits relative to non-employed individuals.

In contrast to these arguments, it is not hard to think of scenarios in which informal employment experience has a negative effect on the prospects of finding a formal job. Barriers to entry to the formal sector may exist that prevent easy transitions from informality.

[1]While there is no agreement on a precise definition to be used in empirical studies, informal jobs comprise a wide range of activities, including small-scale home production for sale, petty trade and untaxed services, self-employment, and wage work that is not formally contracted and not covered by the social safety net (Perry et al., 2007).

[2]Even the most advanced transition economies in Europe have a significant informal sector (Packard et al., 2012). According to Gasparini and Tornarolli (2007), in many Latin American countries the share of informal employment exceeds 50% of the urban labor force. Existing estimates for Sub-Saharan Africa and Asia are even higher (Jütting, Parlevliet, and Xenogiani, 2008). For OECD countries, see Andrews, Sánchez, and Johansson (2011).

In particular, certain labor market institutions like the minimum wage and union collective bargaining agreements might restrict labor demand in the formal sector. Alternatively, low transition rates could result if informality stigmatizes those affected or it carries with it some other kind of "scarring" effect. Prolonged informal sojourns can be associated with losses of the human and social capital that could be required for re-employment in the formal sector (a "lock-in" effect).

For these reasons, the degree of persistence of informality and the extent to which informal jobs are "stepping stones" are important empirical questions. The paper addresses these issues using panel data from Russia, a middle-income country with moderate levels of informality. According to various estimates, informal jobs can account for about 20–25% of employment (Gimpelson and Zudina, 2011, Slonimczyk, 2012). Using simple transition matrices we document persistence rates in the informal sector of almost 50%. The probability of transitioning into a formal sector is also relatively low from the informal sector (about 26%, compared to almost 90% when the origin state is formality). However, because the state of origin is endogenously determined these figures provide a poor indication of the actual degree of state dependence.

In order to disentangle structural state dependence from selection effects we estimate a dynamic multinomial logit model of sector choice that allows for individual heterogeneity in preferences. We apply the method suggested by Heckman (1981) to take care of the initial conditions problem. Using the model to simulate the behavior of individuals in the sample when placed in different counterfactual origin states, we are able to obtain estimates of structural state dependence. As opposed to descriptive transition matrices, our model-based estimates control for a series of observable characteristics. In addition, we use Bayesian inference to obtain an estimate of the position of each individual in the distribution of unobservable heterogeneity. Thus, the estimates of state dependence we present also control, to the extent possible, for unobservable heterogeneity in preferences and ability.

We find that state dependence is much lower than what it would appear based on descriptive evidence. Specifically, for the male sample we find that state dependence is 2.8% for informal employees, 5.7% for formal employees, 13.8% for non-employment, and 20.4% for entrepreneurship. For females, the respective figures are 8.3%, 6.7%, 7,4%, and 20.3%.[3] Importantly, we also find that the chances of transitioning into a formal job are not severely hurt if the origin state is the informal sector. In fact, the likelihood of finding a formal job starting from the informal sector is only a few percentage points lower than when the origin state is a formal job. The simulation results suggest that the specific characteristics and preferences of individuals who occupy the informal sector are the reason behind the relatively high permanence rates and the relatively low transition rates into formal jobs that are observed in simple descriptive transition matrices.

The paper is organized as follows. The next section discusses the issue of mobility in the context of the different theories of the informal sector. Section three introduces the data, sample selection criteria, and gives details on the definition of informality. It also provides a descriptive analysis of mobility across labor market states using transition matrices. Section four describes the empirical model and explains the estimation technique. Section five presents estimation results and evaluates how well the model fits the data. In section six we use the model to simulate different counterfactual scenarios and present our estimates of structural state dependence. The final section discusses our finding and

[3]State dependence (SD) is measured as the average difference between the probability of staying in the same labor market state and the probability of entering from other origin states (see also table 8).

concludes.

2 Informality and Mobility

Theories that explain informality in the labor market can be divided into two competing schools of thought with regards to the issue of segmentation. According to one view, the existence of a large informal sector is explained by rigidities in the urban labor market. For example, in the classic model by Harris and Todaro (1970) a minimum wage set above the market-clearing wage results in rationing of formal jobs.[4] If unemployment benefits are low or nonexistent, workers are left with informal activities as their only option. In other words, there is a perennial excess supply of workers who would want to take formal jobs at the going formal sector wage but are unable to find one. Thus, the formal and the informal sector are segmented.

An alternative perspective sees labor markets as integrated and competitive (Maloney, 1999, 2004). Individuals are endowed with heterogeneous skills, which are valued differently in different sectors (Heckman and Sedlacek, 1985, Magnac, 1991). In addition, jobs vary in non-pecuniary aspects such as amenities and hazards. Individuals choose among the existing employment opportunities in accordance with their preferences and abilities and there are no strict barriers to entry to formal jobs. In more colloquial terms, workers might have good reasons to prefer informal status over formal alternatives due to various desirable characteristics of informal jobs, the low productivity in formal jobs and the poor quality of government insurance programs. This is one of the central messages of the influential World Bank report "Informality: Exit and Exclusion" (Perry et al., 2007).[5]

While segmentation does not preclude the possibility of transitions between formal and informal jobs, it is clear that the two schools of thought have quite different implications regarding the extent of state dependence (are workers trapped in informal jobs?) and the intensity of flows from the informal to the formal sector.[6] If labor markets are segmented flows from formal to informal jobs should be much larger in volume than those going in the opposite direction. Workers are likely to find themselves trapped in the informal sector in presumably inferior (in terms of pay and employment conditions) positions and stay in queue waiting for better (formal) openings. In contrast, the integrated labor markets view implies that there are no real grounds for state dependence in informality and that flows between formal and informal jobs should go in both directions with roughly the same intensity.[7]

Existing studies are inconclusive about which of these theories better represents the reality of labor markets in developing economies. Our knowledge of labor markets in transition countries is particularly shallow.

[4] Calvo (1978) presents a similar argument where the source of rigidities are union-sponsored collective bargaining agreements.

[5] In more recent work, Bosch and Maloney (2010) are less assertive suggesting that while "a substantial part of the informal sector, particularly the self-employed, correspond to voluntary entry, ... informal salary workers may correspond to the standard queuing view."

[6] A third important question involves sectoral wage differentials but it is beyond the scope of this paper.

[7] The two schools of thought are partially reconciled by Gary Fields's idea that the informal sector in developing countries is two-tiered (Fields, 2009). The lower-tier is composed of free-entry jobs and the upper-tier contains skilled jobs. In terms of mobility, the lower tier is expected to be stagnant with one-way entry and with queuing for exit, while the upper tier is integrated with the formal sector. This third position has less clear cut implications regarding the extent of state dependence and the direction of flows between formality and informality.

Gong, Van Soest, and Villagomez (2004) explore mobility patterns using panel data from Mexico. They document relatively large flows across non-employment, formal and informal employment. Using predicted probabilities from a dynamic multinomial model they test the hypothesis that the transition patterns between sectors are in line with the symmetric view of formal and informal sector jobs. These symmetry restrictions on transitions between sectors are not rejected. However, they also find some evidence of entry barriers for low educated individuals.

Pagés and Stampini (2009) provide a comparative study of labor mobility and segmentation in three Latin American countries (Argentina, Mexico, Venezuela) and three transition economies (Albania, Georgia and Ukraine). For all countries they document high mobility rates between formal and informal salary employment but low rates between formal salary and self-employment. For post-socialist countries symptoms of segmentation seem to be clearer than in Latin America. Tansel and Kan (2012) also argue for the existence of segmentation and a static employment structure in Turkey.

Two studies provide evidence on transition countries. Lehmann and Pignatti (2007) analyze flows in the Ukrainian labor market using panel data for the period 2003–2004. They conclude that there is evidence of both dynamism and segmentation, as argued by Fields (1990). In a more recent paper, Lehmann, Razzolini, and Zaiceva (2012) focus on whether displaced workers and voluntary quitters in Russia are more exposed to informality than new labor market entrants or incumbents. They find that displacement entraps workers in involuntary informal employment. Quitters, in turn, experience voluntary informality for the most part, but a minority of them end up in involuntary informal jobs too. The lock-in effect is stronger for workers with low human capital and for those who separate from informal jobs. The latter result also implies that informal employment is persistent. However, the fraction of those involuntary separated in the sample is quite low, so the results are not conclusive.

3 Data

The source of the data for this study is the Russian Longitudinal Monitoring Survey (RLMS). The RLMS is a household panel survey based on the first national probability sample drawn in the Russian Federation.[8] We use data from rounds XI–XX covering the period 2002–2011. These individuals reside in 32 oblasts (regions) and 7 federal districts of the Russian Federation. A series of questions about the household (referred to as the "family questionnaire") are answered by one household member selected as the reference person. In turn, each adult in the household is interviewed individually (the "adult questionnaire").

The structure of the employment module of the adult questionnaire is as follows. First, there are questions about a primary job. Next, individuals can provide information on a secondary job if they have one. Finally, individuals are also asked whether they perform "irregular remunerated activities". The exact phrasing of this last questionnaire item is as follows: "Tell me, please: in the last 30 days did you engage in some additional kind of work for which you were paid or will be paid? Maybe you sewed someone a dress, gave

[8] The RLMS-HSE is conducted by the National Research University Higher School of Economics and the "Demoscope" team in Russia, together with the Carolina Population Center, University of North Carolina at Chapel Hill and the Institute of Sociology RAS. The RLMS website (`www.cpc.unc.edu/projects/rlms-hse`) contains extensive documentation and details on the sampling design.

someone a ride in a car, assisted someone with apartment or car repairs, purchased and delivered food, looked after a sick person, sold purchased food or goods in a market or on the street, or did something else that you were paid for?" The questionnaire structure is such that no one may answer questions on a secondary job unless they have a primary job. However, questions on the irregular activities are independent. In fact, in our sample 7.5% of those considered employed only work doing irregular activities.

The focus of this study is on the main job, defined as the primary job if the individual has one or irregular activities if that is the only source of labor income.

3.1 Sample Selection

The RLMS only started consistently asking questions on informality in 2002. The most recent data are from 2011. Our sample is composed of individuals between 18 and 65 years of age.[9] Since the focus of the study is on mobility we only keep individuals who were observed in at least two consecutive rounds. After dropping a few individuals with missing information on employment status, we are left with an unbalanced panel of 8,547 males and 10,203 females making a total of 42,871 and 53,046 observations respectively. Since mobility patterns are bound to be different across gender lines, we analyze males and females separately.

3.2 Informality Definition

There are two most commonly used definitions of informality: the 'productive' definition and the 'legalistic' or social protection definition. The main difference between them is that while the 'productive' definition focuses on a number of characteristics of the production unit (e.g. the scale of production, whether it is a legal entity independent of the owners, etc.) the 'legalistic' definition focuses on to what extent workers are effectively protected by labor market institutions (e.g. whether social security payments are made). Slonimczyk (2012) discusses in detail the different definitions and how they can be applied using RLMS data. Here we provide only a brief description.

The classification in this paper starts by distinguishing between entrepreneurs and employees at a primary job. The former group is composed of those doing entrepreneurial activities who are either owners of firms or self-employed individuals who work on their own account with or without employees but not at a firm or organization.[10] In principle, it would be possible to distinguish between formal and informal entrepreneurs. As shown in Slonimczyk (2013) the two resulting sub-categories are very small relative to the size of the labor force and have very similar characteristics in terms of hours worked, earnings, turnover rates, and mobility patterns. Since each category considered is computationally (processing time) and statistically (degrees of freedom lost due to extra parameters) costly,

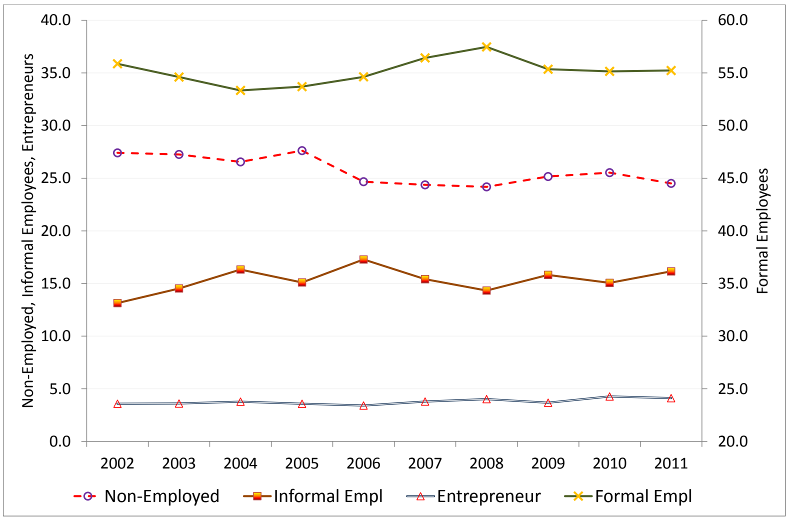
[9]The official retirement age for women is 55 but a large fraction of them keep working until much later. In our empirical model we include controls for age group and pension receipts.

[10]This classification is based on four items of the adult questionnaire: 1) "do you work at an enterprize or organization? We mean any organization or enterprize where more than one person works, no matter if it is private or state-owned. For example, any establishment, factory, firm, collective farm, state farm, farming industry, store, army, government service, or other organization." Enterprize workers are considered entrepreneurs if they answer positively to both 2) "Are you personally an owner or co-owner of the enterprize where you work?" and 3) "In your opinion, are you doing entrepreneurial work at this job?". The distinction between entrepreneurs and employees for non-enterprize individuals is based on: 4) "At this job are you...(a) involved in an employer's or individual labor activity or (b) work for a private individual?"

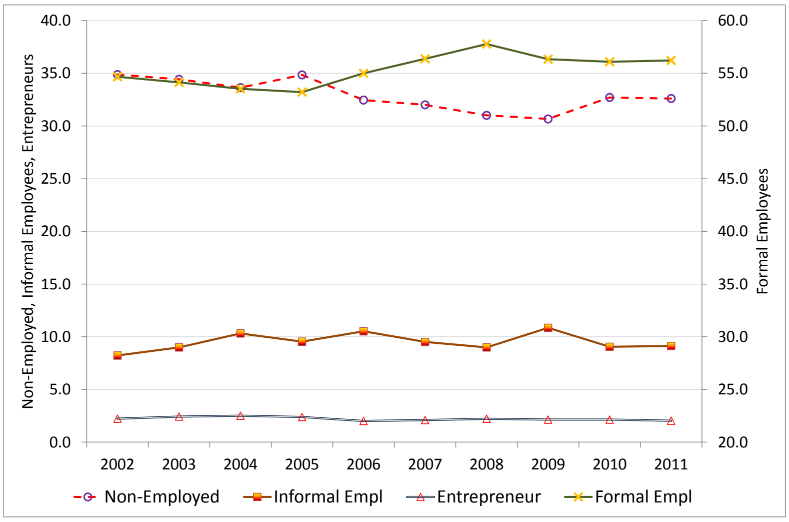
we opted for keeping all entrepreneurs together in one group.

We separate between formal and informal employees as follows. First, following the productive definition, employees not working at firms or organizations are considered informal. Second, for those working at firms or organizations the RLMS questionnaire includes an item that permits determining whether they are registered, i.e. working officially.[11] The Russian labor code mandates that all employees sign a written contract and deposit their 'labor book' with the employer. Therefore, following the social protection criterion, we classify unregistered employees as informal. Finally, individuals without a primary job but who perform irregular activities for pay are also considered informal employees.

Figure 1 – The Evolution of Labor Market Status



(a) Males



(b) Females

Source: Authors' calculations based on RLMS data.

[11]The question is: "Tell me, please: are you employed in this job officially, in other words, by labor book, labor agreement, or contract?"

Table 1 – Transition Matrices: Males

	P-matrix				
	Non-employed	Informal Empl.	Entrepreneur	Formal Empl.	$p_{i\cdot}$
Non-employed	70.9%	15.0%	0.7%	13.3%	25.5%
Informal Empl.	20.4%	49.5%	4.0%	26.2%	15.2%
Entrepreneur	3.3%	15.0%	68.8%	12.9%	3.8%
Formal Empl.	5.5%	6.9%	0.9%	86.7%	55.5%
$p_{\cdot j}$	24.4%	15.7%	3.9%	56.0%	
	T-matrix				
Non-employed		1.09	0.33	1.01	
Informal Empl.	0.99		0.92	1.02	
Entrepreneur	0.38	1.31		1.17	
Formal Empl.	1.04	0.99	0.86		

Notes: Calculations based on 34,324 transitions over the period 2002–2011. The top panel shows the conditional distribution of transitions given the origin state, as well as marginal distributions ($p_{i\cdot}$ and $p_{\cdot j}$). The T-matrix is defined in the main text.

Table 2 – Transition Matrices: Females

	P-matrix				
	Non-employed	Informal Empl.	Entrepreneur	Formal Empl.	$p_{i\cdot}$
Non-employed	78.0%	9.4%	0.4%	12.2%	32.3%
Informal Empl.	25.2%	46.4%	2.5%	25.9%	9.5%
Entrepreneur	5.7%	9.5%	73.6%	11.2%	2.2%
Formal Empl.	6.7%	3.8%	0.5%	89.0%	56.0%
$p_{\cdot j}$	31.5%	9.8%	2.3%	56.5%	
	T-matrix				
Non-employed		1.01	0.35	1.06	
Informal Empl.	1.02		0.99	0.99	
Entrepreneur	0.62	1.23		1.17	
Formal Empl.	1.20	0.79	0.87		

Notes: Calculations based on 42,843 transitions over the period 2002–2011.

Figure 1 shows the distribution of employment status over time. It is important to emphasize the lack of any strong trends in the data. Although we allow for time shocks by including year dummies, the main empirical exercise in the paper assumes the economy is in steady state. Both for females and for males, employment rates have slightly increased over the period. For males, the increase has taken the form of growing informal employment, while entrepreneurship and formal employment have remained roughly constant.[12] Among women, in contrast, the employment gains were in formal jobs, with informality and entrepreneurship constant.

There are visible differences between males and females. First, even though Russian women have relatively high participation rates by international standards, there still exists a gender gap of about seven percentage points. Second, formal employment is relatively more prevalent among women than men. This might reflect the fact that the large public sector is predominantly female. As a fraction of those employed, roughly 74% of males and 83% of females in the sample are formal employees. Correspondingly, there are substantially lower shares of women informally employed and in entrepreneurial roles.

3.3 Labor Market Dynamics

The focus of this study is on individual level mobility in employment status. Tables 1 and 2 present transition matrices for the period under analysis for males and females respectively. The top panel in each table (P-matrix) presents the conditional distribution of labor market destinations given each of the possible states of origin (p_{ij}), as well as the marginal distributions of states of origin and destination ($p_{i\cdot}$ and $p_{\cdot j}$ respectively).

First, note that both for men and women the marginal distributions for origin and

[12] Russia's working age population slightly decreased over the period. See Slonimczyk and Yurko (2013) for a review of the issues and an evaluation of a major pro-natalist policy.

destination states are very similar to each other, which is consistent with the steady state assumption. In fact, the small differences between $p_{i\cdot}$ and $p_{\cdot j}$ are of the same sign as the overall changes shown in figures 1(a) and 1(b) above. For example, the tables show that the proportions of men and women exiting non-employment are slightly higher that the corresponding entry probabilities, leading to a small long-run increase in employment rates. Second, the diagonal elements in the conditional distributions show that non-employment and formal employment are very persistent states. Both among males and females, for example, less than 15% of formal employees leave the state in a given period.[13] In contrast, informal employees appear to be significantly more mobile. A third and final point illustrated by the P-matrices is that, when a change of state takes place, the state of origin seems to affect the likelihood of the possible destinations. For example, men transitioning out of entrepreneurship are much less likely than men leaving informal employment to become jobless.

Analysis based on P-matrices is troubled by the fact that different origin states have very different turnover rates. Also, because the different destination states have different sizes, the conditional distributions are not necessarily informative about the propensities to move from one state to another. Bernabè and Stampini (2009) apply a measure of the propensity to transit from state i to state j that corrects for the turnover rate of the state of origin, as well as for the share of jobs created in each possible destination state. Formally, the elements of their transition matrix are given by

$$t_{ij} = \frac{N_{ij}/(N_{i\cdot} - N_{ii})}{(N_{\cdot j} - N_{jj})/\sum_{k \neq i}(N_{\cdot k} - N_{kk})}$$

where N_{ij} is the number of individuals in state i in $t-1$ and state j in t, and $N_{i\cdot}$ and $N_{\cdot j}$ are the row and column totals respectively.

The lower panels of tables 1 and 2 present the adjusted transition matrices. The results confirm some of the disparities in transition propensities across origin states. For example, both for females and for males it is comparatively harder to become an entrepreneur starting from non-employment relative to other origin states. There are also differences in the ease of access to formal employment but these are relatively small. The T-matrices also show that formal and informal employment are not too different with respect to the risk of non-employment. In fact, the non-employment propensities for formal employees are a little higher than those for informally employed individuals, males and females. Entrepreneurs, in contrast, are very unlikely to become job-less.

While transition matrices, adjusted or otherwise, offer interesting descriptive evidence, they can be misleading because the characteristics of the individuals in the different states are bound to be very different. As an example, suppose there are comparatively more unskilled individuals in non-employment relative to formal and informal employment. If entrepreneurship has high skill requirements, both P- and T-matrices will show relatively lower transition propensities to entrepreneurship from non-employment than from the other two origin states even if the propensities were the same conditional on skill level.

Table 3 presents summary statistics for a number of observable characteristics of the individuals in the sample that are likely to affect transition propensities. It is clear from the table that individuals in different states differ widely in their observable characteristics.

[13]Because the tables do not consider job changes within states, the relative persistence of formal employment does not per se imply low levels of overall mobility. Throughout the paper we use the terms "mobility" and "persistence" in a narrow sense as they apply to transitions across labor market states only.

Table 3 – Descriptive Statistics by Labor Market State

	Males				Females			
	Non-Employed	Informal Employees	Entrepreneur	Formal Employees	Non-Employed	Informal Employees	Entrepreneur	Formal Employees
Age Composition								
15–24	44.8%	21.2%	3.7%	9.5%	33.9%	21.5%	3.2%	9.0%
25–34	9.8%	29.2%	29.6%	29.8%	14.4%	25.5%	21.5%	26.7%
35–44	9.5%	22.7%	35.0%	23.9%	9.4%	21.3%	31.3%	25.4%
45–54	12.9%	18.7%	26.1%	23.9%	12.4%	19.6%	33.8%	26.5%
55–65	23.1%	8.1%	5.6%	12.9%	30.0%	12.1%	10.3%	12.4%
Education Completed								
Less than Secondary	32.3%	19.1%	5.4%	8.2%	24.6%	11.8%	3.8%	4.3%
Secondary School	31.9%	28.9%	20.7%	20.1%	30.6%	28.0%	15.3%	14.5%
Vocational School	19.2%	31.8%	19.5%	30.7%	15.1%	24.4%	19.0%	18.1%
Technical School	8.8%	11.2%	22.6%	18.4%	18.9%	23.6%	36.5%	30.9%
University or Higher	7.8%	9.0%	31.9%	22.5%	10.8%	12.3%	25.5%	32.1%
Region								
Moscow & St Petersburg	9.5%	8.5%	8.7%	12.0%	10.5%	8.5%	6.1%	12.2%
North & North Western	6.1%	4.7%	5.8%	7.1%	5.6%	5.9%	4.4%	8.4%
Central & Black-Earth	15.8%	15.8%	18.0%	19.8%	16.3%	14.3%	22.8%	19.7%
Volga	18.3%	19.4%	17.5%	17.0%	17.8%	16.3%	16.6%	17.5%
North Caucasian	19.1%	19.7%	16.4%	10.7%	17.5%	19.3%	14.4%	9.9%
Ural	13.1%	12.3%	15.0%	15.7%	13.0%	14.9%	13.3%	15.6%
Western Siberian	9.6%	10.6%	10.9%	7.8%	9.4%	10.3%	11.5%	8.0%
East Siberia & Farther	8.6%	9.1%	7.8%	9.9%	9.8%	10.5%	10.9%	8.7%
Other Charact.								
Russian National	70.9%	70.6%	70.6%	79.6%	75.0%	75.4%	75.5%	81.9%
Urban Location	65.0%	64.0%	82.5%	75.7%	67.8%	69.4%	80.7%	78.1%
Married	41.0%	65.5%	89.7%	81.1%	51.9%	56.9%	75.1%	66.9%
Pension	30.7%	8.1%	4.6%	9.6%	37.6%	14.1%	10.8%	15.1%

Notes: The number of observations is 42,896 and 53,090 for males and females respectively.

Thus, transition matrices are bound to present a biased picture. The empirical model we present in the next section is meant to address this issue.

In addition to the problem of differences in observable characteristics, transition matrices do not take into consideration heterogeneity in preferences and skills that are unobservable to the researcher. In particular, if individuals with strong preferences favoring stability are relatively very prevalent in one labor market state, estimated transition propensities out of this state will be biased downwards. A related limitation is that transition matrices do not differentiate between transitions corresponding to the same individual in two different periods and transitions corresponding to different individuals. In contrast, the model we present below incorporates an individual heterogeneity term representing unobserved variation in preferences and other individual characteristics. This term is integrated out in the estimation process, so within- and between-individual variation are given different treatment.

4 Methodology

We model flows among four different labor market states:[14] non-employment ($j = 1$), informal employee ($j = 2$), entrepreneur ($j = 3$), and formal employee ($j = 4$). The individual's utility in each state is specified as

$$U_{itj} = X_{it}\beta_j + Z_{i,t-1}\gamma_j + \alpha_{ij} + \eta_{itj}, \quad j = 1, \ldots, 4 \tag{1}$$

where i and t index individuals and time respectively. The X vector represents observable characteristics influencing state-specific utility. These include variables affecting potential earnings in each state —which we proxy with measures of highest completed education and age—, preferences over non-pecuniary characteristics of jobs as determined by marital status and family structure, and shifts in labor demand over time and across regions. Z_{t-1} is a set of binary variables indicating the labor market state chosen in the previous period (non-employment is the omitted category). The lagged state affects utility through multiple channels, including sector-specific human capital that increases potential earnings, costs associated with job search in different sectors, signalling of unobservable ability, etc. We assume the dynamic process is Markov, so the first lag includes all relevant information regarding sector-specific experience.

Non-observable individual heterogeneity in preferences is represented by α, which is assumed constant over time and independent of the observable characteristics of the individual. In this application there are no a priori restrictions on the range of the heterogeneity, so we specify a normal distribution.[15] Finally, η is a time-varying random component to utility that is assumed independent of the other determinants and has an extreme value distribution.

With these assumptions, the model is a particular case of the mixed multinomial logit (MMNL) class. McFadden and Train (2000) show that any discrete choice model derived from random utility maximization has choice probabilities that can be approximated to any degree by a MMNL model. In particular, MMNL models allow for correlation among state-specific utilities through the individual heterogeneity term, so the independence of irrelevant alternatives assumption is not imposed.

[14] The exact definitions are explained in the data section.

[15] An alternative would have been to specify a discrete distribution with a pre-determined number of mass points (Heckman and Singer, 1984). In this case, it proved computationally infeasible.

As it is only possible to identify differential effects across alternatives, the parameters associated with non-employment $(\beta_1, \gamma_1, \alpha_1)$ are set equal to zero. Conditional on X, Z_{t-1}, and α, utility maximizing individuals choose labor market state l with probability

$$P\left(Z_{it} = l \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i\right) = \frac{\exp(X_{it}\beta_l + Z_{i,t-1}\gamma_l + \alpha_{il})}{1 + \sum_{j=2}^{4} \exp(X_{it}\beta_j + Z_{i,t-1}\gamma_j + \alpha_{ij})}$$

where $\boldsymbol{\alpha}_i \equiv (\alpha_{i2}, \alpha_{i3}, \alpha_{i4})$. Since the random shocks to preferences are i.i.d., the probability of a sequence of choices is simply the product of the time-specific probabilities. Specifically, if individual i chooses a sequence $S_i = (j_1, \ldots, j_{T_i})$, we have

$$P\left(S_i \mid \boldsymbol{X}_i, Z_{i0}, \boldsymbol{\alpha}_i\right) = \prod_{t=1}^{T_i} P\left(Z_{it} = j_t \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i\right) \tag{2}$$

where $\boldsymbol{X}_i$ represents the time sequence of observable characteristics. Importantly, the likelihood in equation (2) is conditional on Z_{i0}, the *initial conditions* of the process.

Since the initial conditions are unobservable, in principle they would have to be integrated out of the likelihood together with the individual heterogeneity. Instead, we proceed as suggested in Heckman (1981) and re-specify the probabilities associated with individuals' first observed period as follows:

$$P\left(Z_{i1} = l \mid X_{i1}, \boldsymbol{\psi}_i\right) = \frac{\exp(X_{i1}\pi_l + \psi_{il})}{1 + \sum_{j=2}^{4} \exp(X_{i1}\pi_j + \psi_{ij})}$$

where π_j are first-period-specific parameters and $\boldsymbol{\psi}_i$ is an individual heterogeneity term.[16] Using this approximation, the likelihood can be rewritten without the need to condition on Z_{i0}:

$$P\left(S_i \mid \boldsymbol{X}_i, \boldsymbol{\alpha}_i, \boldsymbol{\psi}_i\right) = P\left(Z_{i1} = j_1 \mid X_{i1}, \boldsymbol{\psi}_i\right) \times \prod_{t=2}^{T_i} P\left(Z_{it} = j_t \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i\right) \tag{3}$$

The individual heterogeneity is assumed to be normally distributed. In order to impose positive-definiteness in the variance-covariance matrices, we use a Cholesky decomposition and parameterize the diagonal elements in log space. Formally,

$$\boldsymbol{\alpha}_i = \boldsymbol{W}\boldsymbol{\varepsilon}_i, \quad \boldsymbol{\psi}_i = \boldsymbol{W_1}\boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_3)$$
$$\boldsymbol{W} = \begin{bmatrix} \mathrm{e}^{\nu_{22}} & 0 & 0 \\ \nu_{32} & \mathrm{e}^{\nu_{33}} & 0 \\ \nu_{42} & \nu_{43} & \mathrm{e}^{\nu_{44}} \end{bmatrix}, \quad \boldsymbol{W_1} = \begin{bmatrix} \mathrm{e}^{\phi_{22}} & 0 & 0 \\ \phi_{32} & \mathrm{e}^{\phi_{33}} & 0 \\ \phi_{42} & \phi_{43} & \mathrm{e}^{\phi_{44}} \end{bmatrix}$$

[16] The re-specification of the first period probabilities arises from a reduced form approximation to the structural equation (1). Heckman (1981) provides evidence based on a Montecarlo experiment showing that the approximation performs well for a dynamic binary choice model (see also Chay and Hyslop, 2001). Gong et al. (2004) apply the same method to a dynamic model with more than two alternatives.

It then follows that $\boldsymbol{\alpha}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{W}\boldsymbol{W'})$ and $\boldsymbol{\psi}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{W_1}\boldsymbol{W_1'})$. The variance-covariance matrices are uniquely determined by the ν and ϕ parameters, which enter the estimation routine completely unrestricted.

The unconditional individual likelihood can be written

$$\mathrm{L}_i(\boldsymbol{\theta}) = \int P(S_i \mid \boldsymbol{X}_i, \boldsymbol{\alpha}_i, \boldsymbol{\psi}_i)\, \mathrm{d}\boldsymbol{\Phi}(\boldsymbol{\varepsilon}) \tag{4}$$

where $\boldsymbol{\theta}$ represents all model parameters and $\boldsymbol{\Phi}(\cdot)$ is the cdf of a three-dimensional standard normal.

We estimate the model via maximum simulated likelihood (MSL), where the difficult integration in equation (4) is replaced by a simple average over simulations obtained by taking random draws from $\boldsymbol{\Phi}(\cdot)$. Formally,

$$\mathrm{SL}_i(\boldsymbol{\theta}) = \frac{1}{R}\sum_{r=1}^{R} P\big(S_i \mid \boldsymbol{X}_i, \boldsymbol{\alpha}_i = \boldsymbol{W}\boldsymbol{\varepsilon}_i^r, \boldsymbol{\psi}_i = \boldsymbol{W_1}\boldsymbol{\varepsilon}_i^r\big) \tag{5}$$

where $\boldsymbol{\varepsilon}_i^r$ is a three-dimensional vector containing draws from a standard normal. The MSL estimator is consistent and asymptotically equivalent to the usual ML estimator if the number of simulations R grows to infinity at a rate higher than the square root of the number of observations (Hajivassiliou and Ruud, 1994, Train, 2009). For this application, we use $R = 30$ simulations per individual.

The objective function of the MSL procedure is the sum of the log of equation (5) over the N individuals in the sample.

$$\mathrm{SLSL}(\boldsymbol{\theta}) = \sum_{i=1}^{N} \log \mathrm{SL}_i(\boldsymbol{\theta})$$

In order to accelerate convergence, the estimation procedure also calculates the score function:

$$\begin{aligned}\mathrm{g}(\boldsymbol{\theta}) = \frac{\mathrm{dSLSL}}{\mathrm{d}\boldsymbol{\theta}} &= \sum_{i=1}^{N} \frac{1}{\mathrm{SL}_i(\boldsymbol{\theta})} \frac{1}{R} \sum_{r=1}^{R} P(S_i \mid \boldsymbol{X}_i, \boldsymbol{\alpha}_i^r, \boldsymbol{\psi}_i^r) \\ &\times \left(\frac{\mathrm{d}\log P(Z_{i1} = j_1 \mid X_{i1}, \boldsymbol{\psi}_i^r)}{\mathrm{d}\boldsymbol{\theta}} + \sum_{t=2}^{T_i} \frac{\mathrm{d}\log P(Z_{it} = j_t \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r)}{\mathrm{d}\boldsymbol{\theta}} \right)\end{aligned}$$

where, for $l = 2, 3, 4$ and $h \leq l$ the relevant derivatives are

$$
\begin{aligned}
\frac{\mathrm{d}\log P\left(Z_{it}=j_t \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r\right)}{\mathrm{d}\beta_l} &= X_{it}\left[j_{tl} - P\left(Z_{it}=l \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r\right)\right] \\
\frac{\mathrm{d}\log P\left(Z_{it}=j_t \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r\right)}{\mathrm{d}\gamma_l} &= Z_{i,t-1}\left[j_{tl} - P\left(Z_{it}=l \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r\right)\right] \\
\frac{\mathrm{d}\log P\left(Z_{it}=j_t \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r\right)}{\mathrm{d}\nu_{lh}} &= \sum_{j=2}^{4}\left[j_{tj} - P\left(Z_{it}=j \mid X_{it}, Z_{i,t-1}, \boldsymbol{\alpha}_i^r\right)\right]\frac{\mathrm{d}\alpha_j}{\mathrm{d}\nu_{lh}} \\
\frac{\mathrm{d}\alpha_j}{\mathrm{d}\nu_{lh}} &= \mathbb{I}(j=l)\left\{\left[\exp(\nu_{lh})\right]^{\mathbb{I}(l=h)}\varepsilon_{ih}^r\right\} \\
\frac{\mathrm{d}\log P\left(Z_{i1}=j_1 \mid X_{i1}, \boldsymbol{\psi}_i^r\right)}{\mathrm{d}\pi_l} &= X_{i1}\left[j_{1l} - P\left(Z_{i1}=l \mid X_{i1}, \boldsymbol{\psi}_i^r\right)\right] \\
\frac{\mathrm{d}\log P\left(Z_{i1}=j_1 \mid X_{i1}, \boldsymbol{\psi}_i^r\right)}{\mathrm{d}\phi_{lh}} &= \sum_{j=2}^{4}\left[j_{1j} - P\left(Z_{i1}=j \mid X_{i1}, \boldsymbol{\psi}_i^r\right)\right]\frac{\mathrm{d}\psi_j}{\mathrm{d}\phi_{lh}} \\
\frac{\mathrm{d}\psi_j}{\mathrm{d}\phi_{lh}} &= \mathbb{I}(j=l)\left\{\left[\exp(\phi_{lh})\right]^{\mathbb{I}(l=h)}\varepsilon_{ih}^r\right\}
\end{aligned}
$$

In practice, convergence to the optimum required iterating between analytical and numerical derivatives. The estimating routine was written in MATLAB based on code by Kenneth Train.[17]

5 Estimation Results

In this section we present the estimation results for the dynamic multinomial logit model. We also explore how well the model fits the data.

5.1 Parameter Estimates

Table 4 presents maximum simulated likelihood estimates of the coefficients of the dynamic multinomial logit model. The table also presents estimates for the variance-covariance matrix of the individual heterogeneity. Estimates for the initial conditions equation are in table A.1 in the appendix.

The coefficients capture the effect of the independent variables on the probability of choosing each of the employment alternatives relative to joblessness. However, because the model is nonlinear it is difficult to interpret the magnitude of the effects. In the next section we use simulations to get a better sense of the economic significance of some of the factors affecting choices. Here we focus on some salient qualitative results.

First, we find that all the coefficients corresponding to the previous state are positive and statistically significant. This result is unsurprising as one would expect that any form of employment increases the probability of having a job in the next period. The coefficients in the main diagonal are directly related to the extent of state dependence. Both for females and for males, we find that entrepreneurship and formal employment are the states which more strongly attach workers. There is no clear indication that male informal employees are more likely to stay in that state vis a vis other forms of employment. Comparing our model with a dynamic multinomial logit without individual heterogeneity, we find that in

[17]The revised code is available upon request from the authors.

almost all cases the diagonal coefficients are substantially higher in the latter (on average 22% higher).[18] We interpret this as evidence that the individual heterogeneity is removing at least part of the spurious state dependence.

The model incorporates controls for age group. For men, employment probabilities are highest in the 25 to 34 year old category and then decrease with age. The most senior individuals are less likely to be informal employees or entrepreneurs than the baseline group. The pattern for women is somewhat different, with employment probabilities peaking later in life and never quite decreasing to the same level as for 18 to 24 year olds.

The highest completed education level also has a strong impact on employment type. Formal employment and entrepreneurship become more likely as schooling increases. The effect on informal employment seems to be nonlinear, with a university degree decreasing the probability of entering this state relative those with a vocational or technical degree.

We find that ethnic Russians are more likely to get a formal job. Interestingly, Russian women are less likely than women from other nationalities to become entrepreneurs. The main difference across gender lines involves the role of marriage. Married men are more likely to be employed, specially in formal occupations and entrepreneurship. In general, we find the opposite is true for women (the coefficient in the entrepreneurship equation is still positive and significant but small). The number of children in the household has a reinforcing effect for males, leading to even higher employment probabilities. For women, the coefficients are also positive but smaller in size.

Finally, the model picks up some geographic differences. For example, formal employment and entrepreneurship are more likely in cities than in rural areas and less likely in the North Caucasus than in Moscow or St Petersburg. None of the year dummy are statistically significant.

5.2 Model Fit

How well does the model fit the data? Table 5 compares data on individuals' choices to predictions based on the model. The upper panel presents choices for the initial period and an average of the choices for the other periods. In both cases, the model does a remarkable job at predicting the average behavior of the sample.

The lower panel presents transitions disaggregated by origin state. Note that, in contrast to the transition matrices above here we present joint probabilities. While overall the model does a reasonable job tracking the data, there are some small misalignments. The only clear patter is that the model tends to slightly over-predict transitions to formal employment. However, it also slightly under-predicts the fraction of formal employees who do not stay in the state.

[18] The only exception is the entrepreneurship diagonal coefficient in the male sample, which is 1% higher in the model with individual heterogeneity. Estimates for the model without individual effects are omitted to save space but are available from the authors upon request.

Table 4 – Dynamic Multinomial Logit Estimates

	Males			Females		
	Inf Empl.	Entrepr.	Formal Empl.	Inf Empl.	Entrepr.	Formal Empl.
Previous State						
Inf Employee	1.46***	2.11***	1.40***	1.81***	1.84***	1.6***
Entrepreneur	2.32***	6.14***	2.22***	1.63***	6.30***	2.1***
Formal Employee	1.17***	1.74***	2.86***	1.24***	2.04***	3.3***
Age Group						
25–34	0.80***	1.24***	0.96***	0.52***	0.84***	0.57***
35–44	0.41***	0.88***	0.50***	0.71***	1.16***	0.88***
45–54	-0.02	0.38*	0.40***	0.53***	1.11***	0.89***
55–65	-0.60***	-0.46*	0.03	0.09	0.74**	0.23**
Education						
Secondary Compl.	0.04	0.46***	0.55***	0.28***	0.13	0.65***
Vocational School	0.56***	0.70***	1.23***	0.68***	0.50**	1.28***
Technical School	0.45***	1.47***	1.71***	0.62***	0.86***	1.78***
University or more	0.27***	1.84***	2.03***	0.47***	1.16***	2.28***
Other Characteristics						
Russian	0.02	0.09	0.54***	-0.06	-0.36***	0.30***
Married	0.72***	1.30***	1.15***	-0.25***	0.32**	-0.11**
Receives Pension	-1.46***	-1.86***	-2.24***	-1.25***	-1.78***	-1.63***
Size Household	-0.04**	-0.14***	-0.12***	-0.09***	-0.23***	-0.10***
# of Children	0.10***	0.34***	0.24***	0.02	0.27***	0.08**
Urban Area	-0.02	0.76***	0.45***	0.09	0.45***	0.26***
Region						
North & North Western	-0.22	0.08	0.21	0.18	-0.02	0.66***
Central & Black-Earth	-0.03	-0.01	-0.13	0.11	0.23	0.28***
Volga	0.02	-0.27	-0.52***	0.11	0.00	0.03
North Caucasian	-0.08	-0.23	-0.83***	0.32***	0.01	-0.31***
Urals	0.00	0.25	0.04	0.34***	0.20	0.34***
West Siberia	0.12	0.06	-0.55***	0.27**	-0.02	-0.13
East Siberia	0.05	-0.19	-0.15	0.28**	0.22	0.04
Year						
2003	-0.33	-0.66	-0.25	-0.30	-0.41	-0.22
2004	-0.14	-0.65	-0.33	-0.24	-0.70	-0.33
2005	-0.36	-0.92	-0.40	-0.42	-0.74	-0.47
2006	0.04	-0.59	-0.15	-0.02	-0.52	-0.17
2007	-0.15	-0.45	-0.05	-0.28	-0.64	-0.20
2008	-0.13	-0.31	-0.05	-0.25	-0.43	-0.08
2009	-0.19	-0.90	-0.45	0.01	-0.54	-0.24
2010	-0.16	-0.52	-0.36	-0.27	-0.69	-0.41
2011	-0.08	-0.54	-0.29	-0.34	-0.86	-0.41
Constant	-1.49	-5.55	-2.34	-2.10	-5.50	-2.52
Variance-covariance						
Inf Employee	0.6954***			1.1058***		
Entrepreneur	-0.0218	0.0042		1.471***	1.9661***	
Formal Employee	-0.0521	-0.0759***	1.7366***	-0.0477	0.0495	1.3899***
Individuals		8,547			10,203	
Observations		42,871			53,046	
Log Likelihood		-29,076.8			-31,537.2	

Notes: Estimates for the initial conditions are in the appendix. Baseline categories are ‘Not Employed’, ‘18–24 years old’, ‘No Degree’, ‘Moscow-St Petersburg’, and ‘2002’. Significance levels: *** 1%, ** 5%, * 10%.

Table 5 – Model Fit

	Males				Females			
	Nonempl.	Inf Empl.	Entrepr.	Formal Empl.	Nonempl.	Inf Empl.	Entrepr.	Formal Empl.
Choice Probability								
Initial Period	30.5%	14.0%	3.5%	52.1%	38.5%	8.4%	1.8%	51.3%
	30.3%	14.1%	3.5%	52.1%	38.6%	8.6%	2.0%	50.8%
Other Periods	23.1%	15.8%	4.2%	56.9%	29.9%	9.8%	2.4%	57.9%
	23.0%	16.1%	3.6%	57.4%	30.3%	9.9%	1.8%	58.1%
Transition Probability								
Nonemployed	16.6%	3.7%	0.2%	3.1%	23.4%	2.8%	0.1%	3.6%
	14.7%	3.6%	0.1%	5.0%	21.1%	3.0%	0.1%	5.8%
Informal Employees	3.2%	7.9%	0.6%	3.8%	2.4%	4.7%	0.2%	2.4%
	3.1%	5.1%	0.5%	6.8%	2.7%	2.6%	0.1%	4.3%
Entrepreneurs	0.1%	0.6%	3.0%	0.5%	0.2%	0.2%	1.8%	0.3%
	0.1%	0.6%	2.1%	1.3%	0.3%	0.2%	1.0%	1.0%
Formal Employees	3.2%	3.7%	0.5%	49.5%	3.9%	2.1%	0.3%	51.6%
	5.0%	6.8%	0.8%	44.2%	6.2%	4.1%	0.6%	47.1%

Note: White cells contain actual probabilities (data). Gray cells contain model predictions. Transition probabilities are unconditional.

6 Simulations

In this section we use the model to explore the effect of individual characteristics on sector choice. We also analyze the issue of state dependence, i.e. to what extent individuals in one state are bound to stay there.

6.1 The Effect of Observable and Unobservable Characteristics

Education and Age

In order to get a better idea of the economic significance of the effect of observable characteristics, we run simulation exercises in which all individuals in the sample were assigned a counterfactual age or education.[19] Specifically, we assign a fixed value of the characteristic under study while keeping other observables unmodified. Table 6 presents the results.

The first and the second panel in the table are obtained from simulations in which individuals are assumed to have a secondary degree and a university degree respectively. Both for females and for males, higher education levels lead to a significant increase in formal employment. Specifically, having a college degree leads to a 16 (19.2) percentage point increase in the fraction of men (women) with a formal job. Interestingly, education has a very small effect on the fraction of entrepreneurs. The increase in formal employment is explained both by lower levels of nonparticipation and informality. The transition probabilities show that higher levels of education lead to higher retention rates in formal employment and higher exit rates from informality and non-employment. In particular, the probability that a nonemployed individual finds a formal job more than doubles.

We also investigated the effect of age. Employment rates increase at the beginning of the life cycle and then decrease. Interestingly, there is almost no effect across the distribution of employment types. This finding can be interpreted as evidence against the existence of queuing for formal jobs (see also Gong et al., 2004, who find similar results in their study of the Mexican labor market).

Unobservables

The model allows for heterogeneity in preferences and other unobservable determinants of sector choice. How important are these factors vis a vis observable characteristics? In table 7 we present results from simulations in which we have assigned to all individuals an heterogeneity vector with value equal to plus or minus one standard deviation in one dimension and a value of zero in all other dimensions.

Predictably, assigning individuals unobservable heterogeneity value in the formality dimension leads to an increase in the fraction who choose formal employment. What seems remarkable in these simulations is the size of the effect. The gap in the fraction of formal employees with a positive one standard deviation shock and those with a negative shock is 45 and 33 percentage points for men and women respectively. The unobservable component also affects the transition matrices. For example, the probability that an informal employee finds a formal job is 62 and 53 percentage points higher for males and females respectively. These are very strong effects when compared to the findings for education and age.

It should be noted, however, that choice probabilities are not as sensitive with respect to other dimensions of the unobservable heterogeneity. Table 7 also presents simulations

[19]Simulation results for other characteristics are omitted to save space.

Table 6 – Simulating the Effect of Observable Characteristics

	Males				Females			
Secondary Complete								
	NE	IE	ENT	FE	NE	IE	ENT	FE
NE	67.1%	14.6%	0.5%	17.8%	76.0%	10.0%	0.4%	13.7%
IE	23.9%	33.3%	3.3%	39.5%	34.3%	29.4%	1.1%	35.3%
ENT	4.1%	17.9%	49.4%	28.6%	17.1%	9.5%	40.8%	32.6%
FE	12.8%	14.1%	1.5%	71.6%	18.1%	10.0%	1.0%	70.9%
All	27.0%	17.3%	3.5%	52.2%	37.0%	11.9%	1.8%	49.4%
University Complete								
NE	49.8%	12.5%	1.2%	36.5%	56.5%	8.4%	0.7%	34.4%
IE	12.9%	21.6%	5.9%	59.7%	18.0%	18.7%	1.6%	61.7%
ENT	1.4%	7.5%	57.0%	34.2%	6.9%	4.3%	40.9%	48.0%
FE	5.3%	6.7%	1.9%	86.1%	6.4%	4.2%	0.9%	88.5%
All	16.8%	10.4%	4.7%	68.2%	22.5%	6.9%	1.9%	68.8%
Young (18-24 years old)								
NE	64.8%	15.8%	0.5%	18.9%	76.3%	8.5%	0.3%	15.0%
IE	24.7%	33.5%	2.4%	39.4%	36.1%	25.4%	0.8%	37.7%
ENT	3.9%	17.4%	45.4%	33.4%	18.2%	8.9%	32.7%	40.2%
FE	11.2%	13.5%	1.2%	74.1%	15.9%	7.3%	0.6%	76.2%
All	25.7%	17.3%	3.0%	54.0%	36.0%	9.5%	1.3%	53.3%
Midage (35-44 years old)								
NE	56.3%	19.3%	0.9%	23.4%	62.9%	12.7%	0.6%	23.8%
IE	18.2%	35.2%	4.0%	42.7%	22.8%	29.1%	1.3%	46.8%
ENT	2.1%	14.0%	54.5%	29.5%	9.0%	7.6%	41.4%	41.9%
FE	7.8%	13.5%	1.8%	76.9%	8.6%	7.4%	1.0%	83.1%
All	20.7%	18.2%	4.1%	57.0%	26.2%	11.1%	1.9%	60.8%
Senior (55-65 years old)								
NE	69.4%	9.9%	0.3%	20.4%	73.6%	8.8%	0.5%	17.2%
IE	29.4%	23.8%	1.9%	44.9%	33.0%	24.9%	1.4%	40.7%
ENT	5.3%	13.5%	39.6%	41.7%	13.6%	6.4%	43.0%	36.9%
FE	12.3%	8.6%	0.8%	78.2%	13.7%	6.7%	1.1%	78.4%
All	28.2%	11.5%	2.5%	57.9%	33.5%	9.1%	2.0%	55.4%

Notes: Simulated choices when all individuals in the sample are assigned counterfactual characteristics. NE= Non-employed, IE= Informal Employee, ENT= Entrepreneur, FE= Formal Employee.

Table 7 – Simulating the Effect of Unobservable Heterogeneity

	Males				Females			
High FE Component								
	NE	IE	ENT	FE	NE	IE	ENT	FE
NE	45.4%	7.9%	0.3%	46.4%	57.2%	5.1%	0.2%	37.6%
IE	9.9%	12.5%	1.0%	76.6%	15.6%	11.5%	0.4%	72.6%
ENT	1.4%	5.9%	25.1%	67.7%	5.7%	3.9%	19.4%	71.1%
FE	2.0%	1.7%	0.2%	96.2%	2.9%	1.1%	0.1%	95.9%
All	13.5%	5.0%	1.4%	80.1%	20.5%	3.4%	0.6%	75.6%
Low FE Component								
NE	77.1%	17.3%	0.9%	4.7%	85.2%	9.1%	0.3%	5.4%
IE	30.1%	49.4%	5.9%	14.6%	41.6%	37.4%	1.4%	19.6%
ENT	3.6%	16.4%	73.2%	6.9%	14.8%	11.2%	58.8%	15.3%
FE	19.1%	23.2%	3.5%	54.2%	21.7%	10.2%	1.1%	67.0%
All	33.9%	25.5%	6.1%	34.5%	42.5%	12.5%	2.3%	42.7%
High ENT Component								
NE	67.4%	13.8%	0.7%	18.1%	74.3%	7.4%	2.6%	15.8%
IE	21.4%	31.7%	3.5%	43.4%	27.7%	22.7%	8.3%	41.3%
ENT	2.7%	12.2%	57.8%	27.3%	2.6%	1.7%	86.9%	8.8%
FE	7.6%	7.6%	1.0%	83.8%	8.7%	3.6%	3.6%	84.1%
All	23.7%	13.0%	3.7%	59.7%	30.1%	6.5%	5.7%	57.7%
Low ENT Component								
NE	67.4%	13.9%	0.6%	18.2%	75.9%	7.6%	0.0%	16.4%
IE	21.5%	31.9%	3.0%	43.7%	29.9%	24.8%	0.1%	45.3%
ENT	2.9%	13.2%	54.1%	29.8%	16.6%	12.8%	6.6%	64.0%
FE	7.6%	7.6%	0.9%	84.0%	9.0%	3.7%	0.0%	87.2%
All	23.7%	13.0%	3.3%	59.9%	31.3%	7.1%	0.2%	61.4%

Notes: Simulated choices when all individuals in the sample are assigned counterfactual individual shocks. NE= Non-employed, IE= Informal Employee, ENT= Entrepreneur, FE= Formal Employee.

Table 8 – Simulating Dynamics

	Males				Females			
Panel A: Prior Heterogeneity Distribution								
	NE	IE	ENT	FE	NE	IE	ENT	FE
NE	42.7%	18.6%	1.4%	37.3%	57.6%	11.5%	0.7%	30.2%
IE	20.8%	28.5%	3.8%	46.9%	28.0%	24.0%	1.3%	46.7%
ENT	7.7%	19.6%	41.1%	31.6%	16.9%	8.4%	35.6%	39.1%
FE	13.3%	13.4%	1.3%	72.0%	15.2%	7.9%	0.9%	76.0%
SD	28.8%	11.3%	39.0%	33.4%	37.5%	14.7%	34.7%	37.3%
Panel B: Mean of the Posterior Distribution								
NE	25.2%	30.1%	2.4%	42.3%	30.1%	16.8%	15.4%	37.6%
IE	15.5%	32.7%	7.9%	43.9%	24.1%	19.8%	15.8%	40.3%
ENT	3.2%	29.3%	25.9%	41.6%	22.0%	3.2%	36.7%	38.1%
FE	15.4%	30.0%	6.2%	48.4%	22.1%	14.5%	18.0%	45.4%
SD	13.8%	2.8%	20.4%	5.7%	7.4%	8.3%	20.3%	6.7%

Notes: Simulated choices when all individuals in the sample are assigned counterfactual previous state. Panel A assigns a normal distributed heterogeneity term. Panel B assigns the Bayesian posterior mean given the observed choices. NE= Non-employed, IE= Informal Employee, ENT= Entrepreneur, FE= Formal Employee, SD= State Dependence.

when we introduce a plus and minus one standard deviation shock in the entrepreneurship component. The impact of this shock on choice probabilities is negligible for men. For women, the shock to preferences leads to differences in the fraction of entrepreneurs. But even this effect is an order of magnitude smaller than what we find when shocking the formality dimension.

In sum, the simulation exercises presented in this section tell us that individual characteristics have an economically significant effect on sector choice. Estimates that ignore these sources of heterogeneity can be misleading.

6.2 State Dependence

A key question traversing the informality literature is to what extent informal employment is an absorbing state without open exit to formal positions. In order to investigate this issue, we have conducted simulation experiments in which every individual in the sample is assigned a counterfactual previous state. Table 8 presents the results. In the top panel, labeled ‘prior heterogeneity distribution’, we obtained the simulated response by integrating the conditional likelihood over the distribution of the unobservables. These results should be compared against the empirical transition rates (P-matrices) in tables 1 and 2.

The most remarkable feature of the simulated transition matrices is the substantial reduction —relative to the empirical counterparts— in the fraction of individuals who choose to remain in their sector of origin. For example, roughly 50% of informal employees in the data do not change sectors. The simulation results suggest that this statistic is severely inflated by the peculiar observable characteristics of these individuals. Specifically, the model predicts that only 28.5% of males and 24% of females would remain informal after one period if the characteristics of informal workers corresponded to that of the overall sample. Similar reductions are observed for other labor market states.

Uhlendorff (2006) suggests measuring state dependence by the average difference between the probability of staying in the state of origin and the probability of arriving at the state from the other possible origins. We report this statistic in the last row of the panel. According to this measure, both for males and for females state dependence is weakest for informal employment. For males, we find that state dependence is strongest among entrepreneurs and formal employees. We also find that, with the only exception of entrepreneurship, state dependence tends to be higher for females than for males. In particular, non-employment appears to have stronger retention rates among women. In sum, the simulation results confirm that a naive measure of state dependence based on empirical transition rates would be misleading. For example, using the empirical P-matrices would lead to estimates of state dependence of informal employment of 37.2% and 38.8% for males and females respectively. These figures are roughly three times as large as those we found using model simulations.

The reduction in measured state dependence found using simulations would not be too significant if it were compensated with an increase in transitions to undesirable states. Specifically, it would be problematic if the decrease in the fraction of employees that remain informal was explained by an increase in the transitions to non-employment (and viceversa for the lower fraction that remain non-employed). However, this is far from being the case. The simulations show that, both for non-employed and informal employees, the transition rates to formal employment absorb most of the decrease in state dependence. Finally, note that the model symmetrically predicts that, after adjusting for the composition of observables, the transitions from formal employment to non-employment and informal employment would also be larger than what is observed in the empirical P-matrices.

As with the simulation exercises in the previous section, the experiments presented here use the model to answer the question of what individuals in the sample would do under counterfactual circumstances. In this context, using the prior distribution of unobservables to obtain a simulated response is reasonable. However, one could argue that this procedure does not use all available information in an optimal way. An alternative (Bayesian) approach is to use individuals' past choices and characteristics to obtain an estimate of their place in the heterogeneity distribution, which in turn can be used to get simulation results under counterfactual conditions.

We proceed as follows (see Train, 2009, for details). The mean of $\boldsymbol{\alpha}$ in the subpopulation of people who would choose S_i when their characteristics are $\boldsymbol{X}_i$ is

$$\overline{\boldsymbol{\alpha}_i} = \int \boldsymbol{\alpha} h(\boldsymbol{\alpha} \mid S_i, \boldsymbol{X}_i, \boldsymbol{\theta}) \mathrm{d}\alpha \tag{6}$$

where $h(\cdot)$ is the density of the heterogeneity conditional on choices and characteristics (we refer to it as the posterior distribution of $\boldsymbol{\alpha}$). A simulated counterpart can be obtained as

$$\widetilde{\boldsymbol{\alpha}_i} = \sum_{r=1}^{R} w^r \boldsymbol{\alpha}^r; \qquad w^r = \frac{P\big(S_i \mid \boldsymbol{X}_i, \boldsymbol{\varepsilon}_i^r, \boldsymbol{\theta}\big)}{\sum_{r=1}^{R} P\big(S_i \mid \boldsymbol{X}_i, \boldsymbol{\varepsilon}_i^r, \boldsymbol{\theta}\big)}$$

We ran simulations in which individuals were assigned counterfactual origin states and the heterogeneity term is set equal to the mean of the posterior distribution ($\widetilde{\boldsymbol{\alpha}_i}$). The results are presented in the lower panel of table 8. We interpret these results as correcting for the composition of observable characteristics *and also*, to the extent possible, for the position of the individual in the heterogeneity distribution.

In most cases, the extra adjustment leads to further reductions in the diagonal elements of the transition matrix (the exceptions are informal employment for males and entrepreneurship for females, where we find small increases). However, the most noteworthy change involves the trend towards equalization of the conditional probabilities. For example, whereas in the upper panel of the table the probability of landing a formal job for males ranges from 31.6% to 72% depending on the state of origin, the corresponding range in the lower panel is 41.6% to 48.3%. As a result, the measure of state dependence uniformly decreases, both for females and for males and for every labor market state. In particular, state dependence for formal and informal employees are in the single digits. For women, state dependence in non-employment is dramatically reduced in comparison with the simulations using the prior distribution. Finally, note that state dependence in entrepreneurship remains high at about 20% for both females and males. One possible explanation for this finding is the existence of significant sunk costs in tools, equipment and other forms of fixed capital that would be lost in other form of employment.

6.3 Discussion

To our knowledge, there are only two studies with which it is possible to compare our findings. Gong, Van Soest, and Villagomez (2004) estimate a similar model using panel data from Mexico. Note, however, that their data has quarterly frequency and spans a relatively short spell of two years (1999–2000). Also, the model they estimate has three sectors (not working, formal, informal) instead of four and includes interactions between the lagged labor market state and a dummy for higher education. Using their reported simulation results (table 9 in their paper), it is possible to obtain state dependence statistics.[20] Among males, state dependence is about 22% for non-employed, 8% for informal workers and 21% for formal employees. Akay and Khamis (2012) estimate a binary response model (formal/informal) using panel data from the Ukraine. Their estimating sample pools males and females. They report state dependence in informal employment of about 7%.

These studies differ in several important details to ours. However, it is reassuring that a relatively low estimate of structural state dependence in informal employment is not a rare finding.

The relatively low persistence of informal employment and the fact that starting from informality does not severely hurt the chances of obtaining a formal job both lend support to the idea that the labor market in Russia is competitive and relatively flexible. In fact, none of the labor market institutions that are generally seen as possible causes of informality are strict enough to cause segmentation in the Russian setting (Gimpelson and Kapeliushnikov, 2013a). First, the minimum wage and public sector pay are set at a low level. Second, trade unions are weak and have little impact on wage setting. Finally, while labor regulations are strict, they are poorly and selectively enforced. All these factors contribute to remarkable wage flexibility in the labor market, which is hardly compatible with strict dualism or segmentation.

However, a few cautionary remarks are in order. First, not all dualistic theories of the labor market imply strong state dependence of the informal state and low transition rates from informal to formal jobs. There are models in which informality is in a gray area between segmentation and integration. Our findings have little import to these models. Second, our simulation results are based on some difficult to verify assumptions. In particular, we assume the distribution of the individual heterogeneity is normal. A less parametric approach to estimation of the model might lead to different results. Other

[20]The simulations they perform are slightly different from ours. Rather than averaging the responses over the whole sample, they set observable characteristics to benchmark values and the individual heterogeneity terms to zero. They also present separate simulations for high and low skill workers. We focus on the results for males, for whom the calculated state dependence does not vary much by skill level.

assumptions include the markovian structure of the model dynamics and the absence of general equilibrium effects.

7 Conclusion

Informal employment is a serious issue affecting most countries with underdeveloped or weak institutions. In this paper we address the issue of whether informality is a persistent state in which workers are trapped. We also investigate the related question of to what extent transitions from informal to formal jobs are possible. We specify a dynamic multinomial logit model of sector choice that allows for unobservable individual heterogeneity. We use econometric techniques to address the issue of endogenous initial conditions and the computational challenge of integrating out the individual heterogeneity term. The key simulation results are based on estimates of the place occupied by individuals in the distribution of unobservables that are obtained by applying Bayesian inference.

The results provide strong evidence that the mobility patterns observed in empirical transition matrices can be seriously misleading. First, empirical transition matrices ignore the role of education, age, and other observable characteristics in the selection of sector of employment. Model estimates suggest that several of these characteristics have a statistically significant effect on sector choice. Second, neither do the transition matrices account for preferences and other unobservables. Finally, transition matrices do not take into account the panel structure of the data.

The outcome of these biases is that the role of the sector of origin is severely inflated by descriptive evidence. Our simulations show that state dependence is an order of magnitude lower than what P-matrices would imply. A more general point is that the distribution of destination states is much less dependent on the state of origin than it might seem. In sum, the choice of whether and in what sector to work has more to do with fundamentals (preferences, endowments and technology) and less to do with history than what prima facie evidence suggests.

The findings presented in the paper contradict the predictions based on strictly dualistic views of the labor market and lend support to the integrated labor market paradigm. From a policy perspective, the main implication is that the risk of informality is widely spread in the population. Both formal and informal jobs are heterogeneous enough to combine various amenities and costs over which individuals may have differing opinions and preferences.

References

Akay, A. and Khamis, M. The Persistence of Informality: Evidence from Panel Data. In Lehmann, H. and Tatsiramos, K., editors, *Informal Employment in Emerging and Transition Economies*, volume 34 of *Research in Labor Economics*, chapter 7, pages 229–256. Emerald, 2012.

Andrews, D.; Sánchez, A. C., and Johansson, Å. Towards a Better Understanding of the Informal Economy. Working Paper 873, OECD Economics Department, 2011.

Bernabè, S. and Stampini, M. Labour Mobility During Transition: Evidence from Georgia. *Economics of Transition*, 17(2):377–409, 2009.

Bosch, M. and Maloney, W. F. Comparative Analysis Of Labor Market Dynamics Using Markov Processes: An Application To Informality. *Labour Economics*, 17(4):621–631, 2010.

Calvo, G. A. Urban Unemployment and Wage Determination in LDC'S: Trade Unions in the Harris-Todaro Model. *International Economic Review*, 19(1):65–81, 1978.

Chay, K. and Hyslop, D. Identification and Estimation of Dynamic Binary Response Models: Empirical Evidence Using Alternative Approaches. Working paper, University of California at Berkeley, Department of Economics, 2001.

Fields, G. Labour Market Modelling And The Urban Informal Sector: Theory And Evidence. In Turnham, D.; Salome, B., and Schwartz, A., editors, *The Informal Sector Revisited*. OECD, Paris, 1990.

Fields, G. Segmented Labor Market Models In Developing Countries. In Ross, D. and Kinkaid, H., editors, *The Oxford Handbook of Philosophy of Economics*, pages 476–510. Oxford University Press, 2009.

Gasparini, L. and Tornarolli, L. Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata. Working Paper 46, CEDLAS, Feb 2007.

Gimpelson, V. and Kapeliushnikov, R. Labor Market Adjustment: Is Russia Different? In Alexeev, M. and Weber, S., editors, *Handbook of the Russian Economy*. Oxford University Press, 2013a.

Gimpelson, V. and Kapeliushnikov, R., editors. *In the Shadow of Regulation: Informality in the Russian Labor Market*. HSE Publishing House, Moscow, 2013b. In Russian. Forthcoming.

Gimpelson, V. and Zudina, A. "Informals" in the Russian Economy: How Many And Who Are They? Working Paper Series "Labour Markets in Transition" 6, Higher School of Economics, 2011.

Gong, X.; Van Soest, A., and Villagomez, E. Mobility in the Urban Labor Market: a Panel Data Analysis for Mexico. *Economic Development and Cultural Change*, 53(1): 1–36, 2004.

Hajivassiliou, V. A. and Ruud, P. A. Classical Estimation Methods for LDV Models Using Simulation. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics*, volume 4, chapter 40, pages 2383–2441. Elsevier, 1994.

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

Heckman, J. The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In Manski, C. and McFadden, D., editors, *Strutural Analysis of Discrete Data with Econometric Applications*, pages 179–95. MIT Press, 1981.

Heckman, J. and Sedlacek, G. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93(6):1077–1125, 1985.

Heckman, J. and Singer, B. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, 52(2):271–320, 1984.

Jütting, J.; Parlevliet, J., and Xenogiani, T. Informal Employment Re-loaded. *IDS Bulletin*, 39(2):28–36, 2008.

Lehmann, H. and Pignatti, N. Informal Employment Relationships and Labor Market Segmentation in Transition Economies: Evidence from Ukraine. Discussion Paper 3269, IZA, 2007.

Lehmann, H.; Razzolini, T., and Zaiceva, A. Job Separations and Informality in the Russian Labor Market. In Lehmann, H. and Tatsiramos, K., editors, *Informal Employment in Emerging and Transition Economies*, volume 34 of *Research in Labor Economics*, chapter 8, pages 257–290. Emerald, 2012.

Magnac, T. Segmented or Competitive Labor Markets. *Econometrica*, 59(1):165–187, 1991.

Maloney, W. Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico. *The World Bank Economic Review*, 13(2):275–302, 1999.

Maloney, W. Informality Revisited. *World Development*, 32(7):1159–1178, 2004.

McFadden, D. and Train, K. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.

Packard, T.; Koettl, J., and Montenegro, C. *In from the Shadows: Integrating Europe's Informal Labor*. World Bank, Washington DC, 2012.

Pagés, C. and Stampini, M. No Education, No Good Jobs? Evidence On The Relationship Between Education And Labor Market Segmentation. *Journal of Comparative Economics*, 37:387–401, 2009.

Perry, G.; Maloney, W.; Arias, O.; Fajnzylber, P.; Mason, A., and Saavedra-Chanduvi, J. *Informality: Exit and Exclusion*. The World Bank, Washington DC, 2007.

Slonimczyk, F. The Effect of Taxation on Informal Employment: Evidence from the Russian Flat Tax Reform. In Lehmann, H. and Tatsiramos, K., editors, *Informal Employment in Emerging and Transition Economies*, volume 34 of *Research in Labor Economics*, chapter 2, pages 55–99. Emerald, 2012.

Slonimczyk, F. Informality as a Stepping Stone. Working Paper Series “Labour Markets in Transition” 5, Higher School of Economics, 2013.

Slonimczyk, F. and Yurko, A. Assesing the Impact of the Maternity Capital Policy in Russia Using a Dynamic Model of Fertility and Employment. Working paper, Higher School of Economics, 2013.

Tansel, A. and Kan, E. Labor Mobility Across The Formal/Informal Divide In Turkey: Evidence From Individual Level Data. Discussion Paper 6271, IZA, 2012.

Train, K. *Discrete Choice Methods with Simulation.* Cambridge University Press, 2009.

Uhlendorff, A. From No Pay to Low Pay and Back Again? A Multi-State Model of Low Pay Dynamics. Discussion Paper 2482, IZA, Dec 2006.

A Appendix

Table A.1 – Initial Conditions Estimates

	Males			Females		
	Inf Empl.	Entrepr.	Formal Empl.	Inf Empl.	Entrepr.	Formal Empl.
Age Group						
25–34	0.8031***	-5.5519	-0.5423	0.5174***	-5.5017	-0.8589
35–44	0.4093***	1.2432***	-2.3366	0.7052***	0.8432***	-2.5229
45–54	-0.02	0.8846***	0.9598***	0.5296***	1.1577***	0.572***
55–65	-0.6037***	0.3793*	0.5039***	0.0855	1.1097***	0.8803***
Education						
Secondary Compl.	0.0402	-0.4579*	0.4043***	0.2824***	0.7357**	0.8929***
Vocational School	0.5564***	0.4639***	0.0258	0.677***	0.1278	0.2284**
Technical School	0.4514***	0.6983***	0.5541***	0.6238***	0.4956**	0.6516***
University or more	0.2653***	1.4654***	1.2313***	0.4745***	0.8634***	1.2818***
Other Characteristics						
Russian	0.015	-0.1879	-0.5467***	-0.0556	0.216	-0.1261
Married	0.7204***	0.757***	0.5374***	-0.2472***	0.4538***	0.3039***
Receives Pension	-1.4577***	1.3012***	0.4515***	-1.2452***	0.3177**	0.2623***
Size Household	-0.0438**	-1.8567***	1.1455***	-0.0943***	-1.7796***	-0.1144**
# of Children	0.1008***	-0.1374***	-2.2384***	0.015	-0.2251***	-1.6283***
Urban Area	-0.0236	0.0931	-0.1541	0.0883	-0.3625***	0.044
Region						
North & North Western	-0.223	1.8416***	1.7101***	0.1778	1.1625***	1.7847***
Central & Black-Earth	-0.0272	0.0758	2.0325***	0.1058	-0.0165	2.2807***
Volga	0.0165	-0.0136	0.2084	0.1096	0.2315	0.6556***
North Caucasian	-0.0831	-0.2706	-0.127	0.3242***	0.0046	0.2766***
Urals	0.0013	-0.2261	-0.5192***	0.3362***	0.0061	0.0271
West Siberia	0.1187	0.2491	-0.8286***	0.2722**	0.2008	-0.3134***
East Siberia	0.0537	0.0553	0.0411	0.2818**	-0.0233	0.3376***
Year						
2003	-0.3293	0.3411***	-0.1208***	-0.303	0.2717***	-0.1029***
2004	-0.1426	-0.6564	0.2386***	-0.2357	-0.4106	0.0767**
2005	-0.3601	-0.6468	-0.2524	-0.415	-0.6953	-0.2153
2006	0.0386	-0.9248	-0.3293	-0.0249	-0.7433	-0.3337
2007	-0.1455	-0.5914	-0.4011	-0.2774	-0.5209	-0.4723
2008	-0.1261	-0.4541	-0.1523	-0.2481	-0.6388	-0.173
2009	-0.187	-0.3149	-0.0544	0.0063	-0.4331	-0.2048
2010	-0.1577	-0.9018	-0.0533	-0.2676	-0.5366	-0.0782
Constant	-1.4924	-0.0824	-0.5193	-2.1014	-0.3353	-0.689
Variance-covariance						
Inf Employee	0.5643***			0.7685***		
Entrepreneur	0.2092**	0.0794		1.3441***	2.3575***	
Formal Employee	0.0158	-0.0573*	2.1449***	-0.14	-0.1125	2.6848***

Notes: Baseline categories are 'Not Employed', '18–24 years old', 'No Degree', 'Moscow-St Petersburg', and '2002'. Significance levels: *** 1%, ** 5%, * 10%.