Who Migrates and Why?
Evidence from Italian Administrative Data*

Cristian Bartolucci, Collegio Carlo Alberto
Claudia Villosio, Laboratorio R. Revelli
Mathis Wagner, Boston College

February 2014

Abstract

We use twenty years of Italian administrative panel data, a uniquely rich source of information on internal migration experiences from the poorer South to the wealthy North, to identify the role of unobserved worker characteristics in the selection of migrants and returns to migration. We propose and implement a novel iterative estimation method for a switching regression model with the same worker-specific source of unobserved heterogeneity ("ability") present in the selection and both outcome equations. Estimated returns to ability are lower in the north than in the south of Italy and accordingly migrants tend to be drawn from the lower-end of the ability distribution. Around half the gains to migration are due to higher wages, and the other half due to greater labor market attachment. Differential returns to observable characteristics are far less important. Return migration reinforces the original negative selection of migrants, consistent with migrants facing considerable uncertainty about their income in Northern Italy.

JEL Classification: J61, R23, O15.

*We thank Ainhoa Aparicio, Manuel Arellano, Stephane Bonhomme, George Borjas, Jesús Fernández-Huertas Moraga, Giovanni Mastrobuoni, David McKenzie, Mario Pagliero, Thomas Lemieux and members of audiences at University of Milan, Collegio Carlo Alberto, UCL-CReAM, Boston College, Boston University, University of Milan-Bicocca, European Meeting of the Econometric Society, Meetings of the European Association of Labor Economics, the Norface Conference "Migration: Economic Change, Social Challenge "and the Collegio Carlo Alberto Conference on "Legal, Illegal and Criminal Migration"for helpful comments. Corresponding Author: Mathis Wagner, Department of Economics, Boston College, 333 Maloney Hall, 140 Commonwealth Ave, Chestnut Hill, MA 02467, mathis.wagner@bc.edu.

1 Introduction

The economic impact of migration on source and destination countries or regions ultimately depends on who migrates: the "best and brightest" or the "huddled masses."[1] The Roy model, as applied by Borjas (1987) to understanding migration decisions, illustrates that the question of who migrates is inseparably linked to the question of why people migrate. Addressing these questions is particularly challenging since workers' ability is in large part unobserved.[2] In this paper we identify the importance of ability for selection of migrants and returns to migration combining a uniquely rich dataset that tracks migrants in the source and destination region with a novel iterative estimation method for a switching regression model with unobserved fixed heterogeneity.

The data we use, twenty years of Italian administrative panel data, contain detailed information on internal migration experiences and, crucially for our empirical strategy, contains multiple observations on the same individual in source (the poor south of Italy) and destination region (the wealthy North). Using Italy to study migration through the prism of the Roy model has a number of additional important advantages. Italy is a country with a distinct and long-standing North - South divide, and therefore particularly suited to the binary choice techniques that are also used to study Mexico-US migration.[3] Studying internal migration allows us to focus on the impact of wage and employment differentials on migration without the confounding factors that affect cross-country studies.[4] Italy is also a particularly interesting case to study since there is a long tradition of

[1] See, for example, Docquier and Rapoport (2012) for a recent survey of the changing views of the related literature on the effect of brain drain on source countries.

[2] Accounting for unobservables has been shown important for understanding the selection of migrants (Mattoo, Neagu and Özden, 2008; Fernandez-Huertas, 2011; and McKenzie, Gibson, and Stillman, 2010). A Mincer regression, using the data from this paper, with polynomials of potential experience, occupation, and year fixed effects has an R-squared of 0.33. Including individual fixed effects increases the R-squared to 0.79. See Meghir and Pistaferri (2004) and Lemieux (2006) for recent work on the importance of unobservables for wage determination.

[3] See Borjas, Bronars, and Trejo (1992), Dahl (2002), Kennan and Walker (2011) and Bertoli, Fernandez-Huertas and Ortega (2013) for models of migration with multiple possible destinations.

[4] In particular, (unobservable) migration costs are likely to be less important in our context than for international migration. These costs include language barriers, different legal systems, issues related to the transferability of human capital and qualifications, pensions eligibility, unemployment benefits and

emigration from Southern Italy, which is thought to have contributed to impoverishing the South by depriving it of some of its most talented people (Zamagni, 1998).

We estimate a Roy Model where the labor market outcomes in both locations and the migration decision are functions of observable characteristics, the worker's unobserved productivity enhancing characteristics (which we call "ability") and transitory shocks. As is standard in this kind of model we allow for a correlation between transitory shocks to outcomes and the migration decision. What considerably complicates the identification of our migration model is that we also allow our measure of unobserved ability (the worker fixed effect in the outcome equations) to enter in the selection equation. The estimation of this model presents two main challenges: the same source of fixed heterogeneity is present in the three equations, and the estimation of a non-linear model with fixed effects generates inconsistent estimators due to presence of incidental parameters. To solve the first problem we propose a novel iterative estimation method. We first recover an inconsistent estimate of the individual fixed effects in the outcome equation, include these in the selection equation to estimate the inverse Mills ratio, which is then included as a control function to re-estimate the outcome equation and start a new iteration. We iterate to convergence. Monte Carlo simulations suggest that the convergence of this estimator is monotone and fast. To tackle the incidental parameter problem we correct our estimates applying the panel jackknife bias correction presented in Hahn and Newey (2004).

This paper further contributed to the literature by using panel data with multiple observations for individuals in source and destination regions to analyze the degree of selection and the returns to migration as they vary with the duration of migration experience (due to the self-selection of return migrants and assimilation of migrants in the North). Also, the literature has focused on wage differentials as the primary motivation for migrating. The administrative data used in the paper has accurate measures of weeks

---

other aspects of welfare systems, as well as the monetary costs associated with migrating. Evidence by Arellano and Bover (2002) also suggests that the economic forces governing internal and international migration are similar.

worked in a year, enabling us to assess the importance of both differentials in wages and employment opportunities for migration decisions. Finally, we contribute to a substantial literature assessing the validity of the Roy model in the context of migration.

Our findings highlight the importance of accounting for unobserved worker characteristics and differential employment opportunities to understand both the selection of migrants and the returns to migration. Incorporating these two factors in the migration decisions results in clear support for the Roy model of migrant selection as formulated by Borjas (1987). We find that returns to ability are higher in the South and that lower ability workers are more likely to migrate from South to North. The "best and brightest" are found to be more likely to stay in the South, providing evidence against the conventional wisdom that Southern Italy experiences "brain drain".[5] Crucially, selection is driven by unobserved worker characteristics ("ability") and by changes in weeks employed per year. It is essential to account for these when estimating counterfactual wages for migrants in the South, otherwise estimates are biased toward finding positive selection of migrants.

The literature on international migration has highlighted that the existence of unobserved migration costs likely changes the nature of selection, and makes it difficult to identify the role played by differential returns (Chiquiar and Hanson, 2005). For example, for Mexico-US migration McKenzie and Rapoport (2010) find positive selection in communities with weak migrant networks but negative self-selection in communities with stronger networks (and presumably lower migration costs).[6] In our context, except perhaps for the very poorest, liquidity constraints are unlikely to be very important. Rather, our work highlights a hitherto underexplored driver of migration decisions: differential returns to unobserved characteristics. Since these are also important for international mi-

[5]The literature on brain drain has been mostly focused on the consequences of brain drain in the source country. As ability is generally unobserved, little is known on the empirical relevance of brain drain (Carringtong and Detragiache, 1998). See Becker, Ichino and Peri (2003) for evidence of brain drain from Italy.

[6]Papers that have highlighted wealth constraints for understanding international migration patterns include McKenzie and Rapoport (2007), Borger (2010), and Belot and Hatton (2012). An alternative linear utility specification has been proposed by Grogger and Hanson (2011). Fernandez-Huertas (2013) provides an overview of the proposed mechanisms and tests these for Mexico - US migration.

gration, they would provide a further explanation for why existing studies fail to find stronger evidence of negative selection of Mexican migrants to the US, despite the larger variance of wages in Mexico; and the puzzle posed by Grogger and Hanson (2011) who find that migrants are positively selected on educational attainment from almost every sending country in the world, even those countries with very high levels of income inequality.[7] Our finding gives empirical support to Mattoo, Neagu and Özden (2008) which suggests that international migrants may be negatively selected on unobserved ability, with educational attainment a very imprecise indicator of their skills.

Returns to migration, estimated as the average difference between actual annual wages in the North and counterfactual wages in the South, are always positive when accounting for selection on both observed and unobserved worker characteristics. Wage gains from migration are on average 5, 15 and 21 percent in the first, fifth and tenth year respectively, highlighting the importance of assimilation for understanding these returns. In terms of income - where we account for both differences in wages and employment - they are -7, 8, 33, and 50 percent in the first, second, fifth and tenth year, respectively. Around half the gains from migration (after the first year) are due to higher wages, and the other half due to better labor market attachment. The fact that the income gains due to migration are negative in the first year in part reflect the fact that most migration experiences involve an interruption in employment. The returns to migration are lower for high ability workers since the estimated return to ability in the north of Italy (as estimated from the outcome equation in Southern Italy) is significantly lower than the return to ability in the South.

Our data allows us to observe the migration experience of southern Italians who migrate to the north of Italy and eventually return. We find that return migration is remarkably common: around half of those who migrate to the north of Italy return within

[7]See also Ambrosini and Peri (2012), Caponi (2011), Ibarrarán and Lubotsky (2007), Fernandez-Huertas (2011), Kaestner and Malamud (2013), McKenzie and Rapoport (2010) on Mexico-US migration; and Abramitzky (2008), Borjas (2008) and Tunali (2000) for evidence from Israeli kibbutzim, Puerto Rico and Turkey, respectively. Abramitzky, Boustan and Eriksson (2012) provide evidence on the selection of migrants from Norway to the US during 1880 - 1920.

our sample. Two key hypotheses about return migration are that they reflect uncertainty about returns to migration, and are part of a human capital acquisition strategy.[8] We find support for both hypotheses. Returns to migration are significantly higher for those migrants who never return to the South. In the first year the wage gains from migration are 8 percent for non-returnees and 4 percent for returnees, in the fifth year 17 and 10 percent respectively. The income gains are 7 percent for non–returnees and -13 percent for returnees in the first year; and 52 percent for non-returnees and 4 percent for returnees after five years. The evidence is clearly consistent with the idea that a lot of return migration is the result of a disappointing migration experience. Time spent in Northern Italy also has a positive effect on wages in the Southern Italy for return migrants, especially when measured in terms of income. This provides evidence for the human capital acquisition hypothesis, in particular since we control for the selection of return migrants.[9] Return migration also reinforces the original selection of those migrants who remain in the north of Italy, as predicted by a Roy model with uncertainty about outcomes in the destination region (Borjas and Bratsberg, 1996). We find that male migrants who do not return to the South are on average of much lower ability than those who return, reinforcing negative selection of migrants from the ability distribution for both wages and income.

The remainder of the paper proceeds as follows. Section 2 outlines the theoretical framework we use to think about migration. Section 3 provides background information on Italy, discusses the data and presents some preliminary evidence. We present our empirical strategy in Section 4. The results are discussed in Section 5. Section 6 concludes.

[8] A number of studies, including Borjas and Bratsberg (1996), Dustmann and Weiss (2007), Thom (2010), and Dustmann, Fadlon and Weiss (2011) develop models of temporary migration in which migrants acquire additional skills while working abroad that are rewarded in the home country.

[9] Reinhold and Thom (2011), De Coulon and Piracha (2005), Co, Gang and Yun (2000), Hunga, Barrett and Goggin (2010) find that Mexican, Albanian, Hungarian and Irish, respectively, return migrants command a wage premium. Lacuesta (2006) attributes the gains to Mexican return migrants to selection.

2 Theoretical Framework

We begin by outlining a theoretical framework within which to analyze the questions of who migrates and why. The standard framework for thinking about migration decisions is a version of the Roy model (Roy, 1951), adapted for understanding migration decisions by Borjas (1987). For a discussion of the empirical content of the Roy model see Heckman and Honore (1990).

Consider an individual i who every period t has the choice whether to migrate $M_{it} = 1$ or not $M_{it} = 0$ between a source location $j = s$ and a single destination location $j = n$. We assume she makes that decision based on the difference in outcomes y_{ijt}, typically income or wages, in each location and a one-time migration cost c_{it}. The migration decision is carried out according to

$$M_{it} = \begin{cases} 1 & \text{"migrate" iff } y_{itn} - y_{its} - c_{it} \geq 0 \\ 0 & \text{"stay" iff } y_{itn} - y_{its} - c_{it} < 0 \end{cases}$$

The location-specific outcomes are modeled as a function of time-varying observable characteristics x_{it} and an index of time invariant unobserved characteristics α_i

$$y_{itn} = \rho_n \alpha_i + \beta_n x_{it} + u_{itn}, \tag{1}$$

$$y_{its} = \rho_s \alpha_i + \beta_s x_{it} + u_{its}, \tag{2}$$

where u_{itj} are potentially correlated location-specific transitory shocks,[10] which are distributed independently of α and x. The prices of observed and unobserved worker characteristics are location-specific: the return to ability is given by ρ_j, and the returns to observable characteristics by β_j.[11]

In the basic Roy model, the parameters of the selection equation are a function of the parameters of the outcome equations. However, we estimate the model without imposing

[10] We allow for correlation in shocks across locations. For example a shock to the labor market performance of a worker in the South can be a result of a health shock which would also affect her performance in the North.

[11] These equations could be interpreted in terms of the present value of the earnings stream in each country, a reformulation which would fit within the human capital investment framework proposed by Sjaastad (1962). They could also be expressed in terms of log-linear utility, see for example Dahl (2002).

constrains in the selection equation. We check *ex-post* if the estimated parameters in the selection equation are consistent with the estimates of the outcome equations. This test allows us to further evaluate the fit of the Roy model. Thus we rewrite the selection equation:

$$M_{it} = 1 \left(m \left(\alpha_{i,} x_{it}, z_{it} \right) + v_{it} \geq 0 \right), \tag{3}$$

where z_{it} are individual characteristics that are excluded from the outcome equations and v_{it} are unobserved individual-specific time-varying factors, these are assumed to enter additively separable and distributed independently of α_i, x_{it} , and z_{it} with mean zero and variance σ_v^2. The actually observed outcome y_{it}^* is

$$y_{it}^* = y_{itn} M_{it} + y_{its} \left(1 - M_{it} \right), \tag{4}$$

Following Heckman (1976, 1979), Lee (1976) and Maddala (1983) we reformulate the outcome equations conditional on actually being observed. As is typical and analytically convenient, we assume that the unobservable transitory shocks in the selection and outcome equations are normally distributed. Then

$$y_{itn}|_{M=1} = \rho_n \alpha_i + \beta_n x_{it} + \frac{\sigma_{nv}}{\sigma_v} \frac{\phi \left(m \left(\alpha_i, x_{it}, z_{it} \right) \right)}{\Phi \left(m \left(\alpha_i, x_{it}, z_{it} \right) \right)} + \varepsilon_{itn}, \tag{5}$$

$$y_{its}|_{M=0} = \rho_s \alpha_i + \beta_s x_{it} - \frac{\sigma_{sv}}{\sigma_v} \frac{\phi \left(m \left(\alpha_i, x_{it}, z_{it} \right) \right)}{1 - \Phi \left(m \left(\alpha_i, x_{it}, z_{it} \right) \right)} + \varepsilon_{its}, \tag{6}$$

where σ_{jv} is the covariance between u_{itj} and v_{it} the time-varying idiosyncratic components of the outcome and selection equations, ϕ is the standard normal probability density function, Φ is the standard normal cumulative density function, and ε_{itj} are mean zero residuals which are by construction independent of α_i , x_{it}, z_{it} and v_{it}. The terms $\frac{\phi(m(\alpha_i, x_{it}, z_{it}))}{\Phi(m(\alpha_i, x_{it}, z_{it}))}$ and $-\frac{\phi(m(\alpha, x, z))}{1 - \Phi(m(\alpha, x, z))}$ are known as control functions and are the standard inverse Mills ratios. See the Appendix for an extension of this model to incorporate return migration.

We are now in a position to more precisely characterize what we mean by the selection of migrants and the returns to migration. The literature on the selection of migrants is interested in how, in the source country, the distribution of outcomes for migrants $y_{its}|_{M=1}$ differs from the distribution of outcomes for non-migrants $y_{its}|_{M=0}$. The literature on the returns to migration is interested in the gains experienced by a migrant $y_{itn}|_{M=1} - y_{its}|_{M=1}$, and possibly also the potential gains for non-migrants $y_{itn}|_{M=0} - y_{its}|_{M=0}$. The difficulty of course is that the counterfactual distribution of outcomes in the source country for migrants is not observed (and similarly, the counterfactual distribution of outcomes for non-migrants in the destination country is not observed). The central challenge in estimating the counterfactual outcome distributions, required to characterize the selection of migrants and estimate the returns to migration, is that the selection of migrants can be driven by observable characteristics of migrants x_{it}, but also unobserved characteristics α_i and individual-specific temporary shocks u_{it}.

Borjas (1987, 1991) develops theoretical predictions from a simplified version of this model where migration costs (or location preferences) are assumed to be uncorrelated with observed and unobserved worker characteristics. A key insight is that the variance of log wages reflects the return to skills, with a higher variance implying higher returns. The empirical prediction is that if the variance of log wages is higher in source than destination country migrants will be disproportionately drawn from the lower tail of the source country's skill and wage distribution (negatively selected), i.e. less skilled, lower wage workers are more likely to migrate. If the variance of log wages is higher in destination than source country migrants will be disproportionately drawn from the top end of the source country's skill and wage distribution (positively selected), i.e. high wage individuals are more likely to migrate. If migration costs systematically vary with the skill-level and wage of a worker the nature of selection is affected and these predictions possibly over-turned. For example, Chiquiar and Hanson (2005) suggest that the reason they find that Mexican migrants to the US are selected from the middle of the wage distribution is that migration costs are very high for low-skilled Mexicans.

Potential migrants may be uncertain about the economic conditions they will face after migration, i.e. u_{itn} is not, or not fully, observed. As long as return migration costs are relatively low, workers who experience worse than expected outcomes in the destination region may wish to return to their home. Borjas and Bratsberg (1996) use the Roy model to describe the type of selection that characterizes return migrants. They suggest that the return migration decision reinforces the original type of selection of migrants. Since they are the marginal immigrants, those with the lowest returns to migration, who are most likely to become return migrants, migrants who stay in the destination region are the "best of the best" if there is positive selection and the "worst of the worst" if there is negative selection.

3 Background and Data

3.1 Background

The Italian peninsula has historically been a highly heterogeneous place, frequently invaded and settled by a variety of people, geographically fragmented by the Apennines and linguistically fragmented into frequently mutually incomprehensible dialects. In 1861, the year the Kingdom of Italy was born, it has been estimated that one Italian in forty spoke Italian: just over 630,000 people out of a total of 25 million. Even adding those with some familiarity with the language it is difficult to push the figure beyond 10 percent. In Italy nearly everyone spoke in dialect, not just peasants and artisans and the urban poor, but merchants, aristocrats and even monarchs.[12]

At the time of unification what we consider the south of Italy was all part of the Kingdom of The Two Sicilies (except Sardegna, which was ruled by the Piedmontese), that had been created in 1814 at the Congress of Vienna after the Napoleonic Wars. Economic statistics reveal how separate the kingdom was from the rest of Italy: in 1855 85 percent of its exports were sent to Britain, France and Austria, while only 3 percent

[12] Gilmour (2011) provides an enjoyable read on the diversity from which modern Italy emerged.

crossed the border into the Papal States."The place [Naples] was different, a distinct, cosmopolitan entity, a kingdom (with or without Sicily) with an ancient history and borders which, almost uniquely in Italy, were not subjected to rearrangement after every war" Gilmour (2011, p. 143).

To this day there are huge economic, political and cultural differences between these regions. Southern Italy's GDP per capita is around 60 percent and unemployment rates around double those of Northern Italy (see Figure 1). As a result there has been large-scale and well documented outward migration from Southern Italy, to the North and abroad. While emigration flows peaked before World War I and just after, Southern Italy has continued to experience large outward migration. In the period we are considering for our analysis the migration rate between the South and north of Italy was around 0.45 percent per annum in the 1980s and rose to 0.7 percent in 2000 and continued at around 0.6 percent thereafter, with migration in the other direction at around 0.2 percent (Del Boca and Venturini, 2003).[13] The gross emigration rate out of Italy is around 0.1 percent per annum during this period (Bonifazi et al., 2009).[14] To put these numbers into perspective, as share of Mexico's national population, the number of Mexican immigrants living in the U.S. increased from 3.3 percent in 1980 to 10.2 percent in 2005, an annual net migration rate of somewhat less than 0.3 percent (Hanson and McIntosh, 2010). Though, just as Italy, Mexico-US migration has been distinguished by a high propensity for return migration (Massey, Durand and Malone, 2003).

[13] These statistics are computed from the National Institute for Statistics (Istat) local registers which report the change of residency for the whole population in any given year. The data do not distinguish between workers and family members. From the social security data on the employed population used in this paper we get that on average, in the period considered for our analysis, 13 percent of the workers in the North were born in the South, compared to a 3.3 percent of workers in the South born in the north of Italy.

[14] Emigration rates vary by skill group. For example, Becker, Ichino and Peri (2003) find that in the late nineties, between 3 and 5 percent of the new college graduates from Italy emigrated.

3.2 Data

The data used for this paper is the Work Histories Italian Panel (WHIP), a database of individual work histories randomly selected from all Italian Social Security Administration (INPS) archives. WHIP represents a sample of about 1 percent (sampling ratio 1:90) of all individuals who have worked in Italy from 1985 to 2004. For each of these people their entire working career is observed if they are enrolled in private, self-employment or atypical contracts, but also if they are in retirement spells or in non-working spells in which they receive social benefits (i.e. unemployment subsides or mobility benefits). Individuals who have an autonomous social security fund, namely people who work in the public sector or as free-lancers (lawyers or notaries), are not observed in WHIP.[15]

Southern Italy, *Il Mezzogiorno*, is composed of the regions of Abruzzo, Basilicata, Campania, Calabria, Puglia, Molise, Sicilia and Sardegna. All other regions, Piemonte, Valle D'Aosta, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Liguria, Emilia-Romagna, Toscana, Umbria, Marche, Lazio comprise the center-north of Italy, see Figure 2. In this paper we focus on two outcome variables: wages and income. The average weekly wage for a worker is calculated as the total income earned in a year divided by the full-time equivalent weeks employed.[16] Income is measured as the product of the weekly wage and the weeks employed per year. The average weekly income is the total income earned in that year divided by fifty-two.[17] The difference between these two measures

[15] The data does not capture informal employment. According to Istat estimates in the nineties on average informal employees accounted for about 13 percent of the total employment in Italy. Informality is around 7 percentage points higher in Southern Italy than national average. This share varies considerably across sectors: from more than 40 percent in agriculture, to around 7 percent in manufacturing.

[16] A reason we use the weekly rather than daily wage is that working days are commonly underreported by firms to adjust the total wage bill to the minimum wage requirements. Furthermore, such underreporting does not seem to be distributed uniformly across the country, but it appears to be more frequent in the South and among blue collar workers (Contini, Filippi and Malpede, 2000). Underreporting of weeks worked is likely a lot less severe.

[17] We deflate wages so that the sample mean in every year is identical (at 2004 levels), thereby accounting for both inflation and general productivity growth. We do the same for the income measure, which also removes variations in the average number of weeks employed across years.

In cases where an individual has more than one job in a year, the job characteristics are those associated with the longest employment spell.

Note that migrants who move within a year have an observation in both the south and north of Italy of earnings is that the income measure takes account of both the weekly wage and the average number of weeks employed in a year. The wage is the variable typically used in this literature, since it is unusual to have accurate measures of weeks worked, but it fails to account for the fact that employment opportunities may differ between locations. The income measure, by accounting for both wage and employment differentials, is a more complete measure of an individual's earnings opportunities.[18]

Table 1 presents different measures of wage inequality in the South and North.[19] It shows that the variance of log wages and the Gini coefficient are both higher in Southern Italy (and a little more so if we account for weeks employed per year). This implies that earnings inequality, as it is usually measured, is higher in the south than in the north of Italy. Controlling for observable characteristics the variances of log wages in the North and in the South shrink but the one from the South decreases relatively less than the variance of log wages in the North, and therefore the gap in inequality increases. The implication is that if moving costs are small (or only weakly correlated with worker attributes) we should observe that in Italy workers with lower wages and lower unobserved skills are more likely to migrate from South to North.

All of our analysis is based on those born in Southern Italy and first observed working there, who we define as our pool of potential migrants. We focus on individuals born between 1946 and 1975, so as to ensure that we have a sufficient number of observations for most individuals. We exclude apprentices, training-on-the-job contracts and self-employed from our analysis since we are concerned about how accurately the available

in the same year.

[18]We think of the wage as reflecting workers' productivity and hence not something that workers choose directly (they do choose indirectly of course by, for example, investing in skills and education). In contrast, the number of weeks worked is possibly, though not necessarily, something the worker chooses. The interpretation of the income measure is most straightforward if number of weeks worked is exogenous, determined by job destruction and job findings rates that can not be directly affected by the worker. If weeks worked per year can be directly affected by workers, who might vary their job search intensity for example, then the factor model described in the previous section is at best a reduced-form model describing a more complicated choice process.

[19]To calculate these we impose the same age, birth cohort and type of work restrictions as for our main analysis.

income measure reflects the human capital of these individuals. In addition, we exclude those observations associated with employment that is intrinsically temporary: seasonal workers, fixed-term contracts and temporary workers, which make up 0.6 percent, 2.5 percent and 0.5 percent respectively of all contracts in our data. Finally, we include only those workers between 20 and 50 years old, based on our prior that the very young are much more likely to be tied movers or move due to education rather than labor market related reasons; and that for the old we start seeing a lot of retirement.[20]

Our final sample includes 31,626 unique individuals (22,685 men and 8941 women), 16 percent of whom at some time migrate to the center-north of Italy (19 percent of men, 9 percent of women). Of those who migrate to the North 54 percent return to the south of Italy within our sample (58 percent of men, 34 percent of women). The mean number of observations per individual is 9. A total of 206,324 and 16,536 observations for men in Southern and Northern Italy, respectively, and 60,415 and 2,503 observations for women.

An important issue is the degree to which the administrative data used in this paper is representative of the population of Southern Italy, and thus whether sample selection may bias our findings. Probably the key concern is that we do not observe those who are employed informally employed. With as much as 20 percent of Southern Italy's workforce informally employed that may of course substantially affect our results. Similarly, the exclusion of public sector workers is problematic. To assess the importance of our sample restrictions we use the Italian module of the EU Statistics on Income and Living Conditions data set (IT-SILC). The IT-SILC survey includes a representative sample of all individuals in Southern Italy. In Figure 3, we compare the wage distribution of workers in

[20]We also drop the top and bottom 1% of weekly wage observations to deal with outliers caused by, for example, coding error.

We do not attempt to model the decision of those who do not work at all in a given year, and therefore in our analysis only include those observations where the individual earns a positive income in that year.

Our subsequent analysis is only possible for individuals who we observe at least twice (after all other sample restrictions), hence we exclude those who are observed only once.

Note that we do not capture any migrants who exit the labor force on account of the migration decision, for example, women who move with their husbands and are subsequently out of the labor force in the North.

our sample and those in the IT-SILC, both referring to the whole economy and restricted to the south.[21] We find that the distribution almost overlap, with small differences only in the tails of the distribution. This suggests that, at least in terms of wages, our sample may be representative of the whole population.

The other limitation of the data is that we do not observe emigration out of Italy, which is around 0.1 percent per annum during this period (Bonifazi et al., 2009). The existing evidence suggests that emigration from Southern Italy to other countries, as opposed to the North, is unlikely to substantially affect our results. Becker, Ichino and Peri (2003) find that while emigration from Italy has been increasingly high skilled during the nineties, with between 3 and 5 percent of the new college graduates going abroad each year, this phenomenon is largely restricted to the Northern Italy. Consistent with our findings they cite a tendency of educated workers to stay in the South (see also Gloria and Ichino, 1994).

The variables in the dataset available for an employment spell are the total income earned during that spell, the duration of the spell in weeks, as well as the full-time equivalent number of weeks worked in the spell (accounting for part-time work), the age of the worker, the gender, the place of birth, the type of contract (open-end, fixed term, seasonal worker), an indicator for part-time or full-time employment, the occupation (blue collar, white collar, managerial), sector of economic activity (by 1-digit NACE) and the region of work. We use the data to construct workers' tenure at an establishment and various mobility indicators. The mobility indicators are a worker's average annual job switches (i) within a 1-digit industry, (ii) across 2-digit industries and (iii) across regions.[22] See Table 2 for some descriptive statistics on the sample.

[21] The IT-SILC covers the informally employed but does not record whether they are informally employed. Therefore a direct comparison of the wages of the formally and informally employed is not possible.

[22] One shortcoming of the data is that there is no information about the educational attainment of workers. To evaluate how important the omission of educational attainment is in predicting wages in Italy we use again the IT-SILC dataset , which contains both education and occupation variables. We find that education explains 2.5 percent of the variation of residuals from a wage regression controlling for experience and occupation when using the 2004 cross-sectional sample, and almost zero when using the 2004-2007 longitudinal version of IT-SILC with individual fixed effects. See the Appendix for further

4 Empirical Strategy

In this section we present a novel iterative algorithm that allows us to estimate the full model as described in Section 2. The identification of our migration model is considerably more complicated than the standard selection model (see Maddala, 1983, who gives complete details for this model, which he calls a switching regression model with endogenous switching). This is because we allow the time invariant unobserved worker characteristics to enter the selection and both outcome equations. In the estimation we consider the unobserved fixed characteristics as worker fixed effects, allowing an unrestricted correlation between the worker unobserved fixed characteristics and the observed ones. The estimation of the model presents two main challenges: First, the same source of fixed heterogeneity is present in the three equations. Second, the estimation of a non-linear model with fixed effects generates inconsistent estimators due to presence of incidental parameters.

To solve the first problem we propose a novel iterative estimation method, which extends the standard switching regression model. We parameterize the selection equation (3), as follows

$$M_{it} = \mathbf{1}[m(\alpha_i, x_{it}, z_{it}) \geqslant \upsilon_{it}] = \mathbf{1}(\gamma\alpha_i + \theta_x x_{it} + \theta_z z_{it} \geqslant \upsilon_{it}) \tag{7}$$

The outcome equations are given by (1) and (2). However, we can not identify both ρ^s and ρ^n. Therefore, we normalize the price of α in the South, and we identify the price differential in the North as a loading factor. As described in equations (5) and (6), the model can be written in terms of conditional means. Assuming that υ_{it} is standard normal distributed:

$$E(y_{it}^S | x_{it}, \alpha_i, m_{it} > \upsilon_{it}) = \alpha_i + \beta^S x_{it} - \sigma_{\upsilon\varepsilon}^S \lambda_0(\gamma\alpha_i + \theta_x x_{it} + \theta_z z_{it}) \tag{8}$$

and

information.

$$E(y_{it}^N|x_{it},\alpha_i,m_{it}<v_{it}) = \rho\alpha_i + \beta^N x_{it} + \sigma_{\upsilon\varepsilon}^N \lambda_1[(\gamma\alpha_i + \theta_x x_{it} + \theta_z z_{it})] \quad (9)$$

where $\lambda_0(.) = \phi(.)/\left(1-\Phi(.)\right)$ and $\lambda_1(.) = \phi(.)/\Phi(.)$ are the inverse mills ratios.

In the standard switching regression model the inverse Mills ratio is estimated in a first step by fitting a discrete choice model on the migration decision. In the second step, the outcome equations are estimated using linear regression, by including the inverse Mills ratio as a regressors that adjusts for selection bias. The identification of our model is more complicated since α_i is unobserved. We use the following iterative algorithm:[23]

1. We first calculate an inconsistent estimator $\hat{\beta}_1^S$ of β^S using a within group (individuals) estimator of equation (8);

2. Then use β_1^S to recover an inconsistent measure of the worker specific constant $\hat{\alpha}_{i1}$;

3. We proceed to use $\hat{\alpha}_{i1}$ to calculate $\hat{\theta}_{x1}, \hat{\theta}_{z1}$ and $\hat{\gamma}_1$ by estimating a probit which fits the conditional probability that $m_{it} > v_{it}$ in equation (7);

4. Use $\hat{\theta}_{x1},\ \hat{\theta}_{x1},\ \hat{\gamma}_1$ and $\hat{\alpha_1}$ to calculate the inverse Mills Ratio $\lambda(\hat{\gamma}_1\hat{\alpha}_{i1} + \hat{\theta}_{x1}x_{it} + \hat{\theta}_{z1}z_{it})$;

5. Use $\lambda(\hat{\gamma}_1\hat{\alpha}_{i1} + \hat{\theta}_{x1}x_{it} + \hat{\theta}_{z1}z_{it})$ to calculate $\hat{\beta}_2^S$ using a within group (individuals) estimator of equation (8);

[23] We are not the first to tackle a problem which combines selection bias and unobserved fixed heterogeneity. In the case of the standard Heckman selection model with two equations, Wooldridge (1995) propose a method including two sources of heterogeneity, but this method is only valid to identify β, θ_x and θ_z, but not γ, the correlation between both sources of heterogeneity. However in our case, it is exactly that correlation which is informative about selection on ability. Verbeek and Nijman (1992) and Zabel (1992) consider a random effects model under the assumption of normality and serial independence of the idiosyncratic errors in both the selection and the outcome equations. Although the later method can allow for some correlation between observable characteristics and both sources of unobserved heterogeneity it is more demanding in terms of distributional assumptions. It involves assumptions on the distribution of shocks in both equations and in the distributions of both sources of unobserved heterogeneity. The main problem being, beyond the computational demand involved in its estimation due to the requirement to evaluate multiple integrals, that most of these assumptions are hard to test. Hoderlein and White (2009) suggest a significantly easier estimation procedure with which they are able to recover the coefficient of the observable characteristics, β. However, γ, which is of primary interest in this paper remains unidentified.

6. We keep iterating on steps 2 to 5 until $\begin{pmatrix} \beta^S \\ \sigma_{v\varepsilon} \\ \theta_x \\ \theta_z \\ \gamma \end{pmatrix}_M = \begin{pmatrix} \beta^S \\ \sigma_{v\varepsilon} \\ \theta_x \\ \theta_z \\ \gamma \end{pmatrix}_{M-1}$ where M is the number of iterations.

7. Once we have estimates of $\hat{\beta}^S$, $\hat{\theta}_z$, $\hat{\theta}_x$ and $\hat{\gamma}$, and measures of $\hat{\alpha}_i$, we estimate β^N and ρ by OLS in equation (9), including $\hat{\gamma}$ and the inverse Mills Ratio as regressors.

Monte Carlo simulations suggest that the convergence of this estimator is monotone and remarkably fast (see Appendix).

Our second problem is the presence of incidental parameters. The estimates produced by our method are in general inconsistent for fixed T. Since Newman and Scott (1948), it is well known that treating individual effects as separate parameter to be estimated is typically subject to the incidental parameter problem. In this case the estimation of the parameter of interest will be inconsistent if the number of individuals goes to infinity while the number of time periods is held fixed. The inconsistency is due to the finite number of observations that are used to estimate each individual specific parameter. Therefore, the estimation error for the individual effects does not vanish as the sample grows in the number of individuals.

In order to tackle the incidental parameter problem we correct our estimates applying the panel jackknife bias correction presented in Hahn and Newey (2004). The panel jackknife is an automatic method of bias correction. To describe it let $\hat{\theta}_{(t)}$ be the estimator based on the subsample excluding the observations of the t^{ht} period. The jackknife estimator is:

$$\hat{\theta}_{BC}^{Jackknife} = T\hat{\theta} - (T-1)\sum_{t=1}^{T}\hat{\theta}_{(t)}/T$$

Monte Carlo examples, presented in the Appendix, show that the bias correction substantially reduces the incidental parameter problem. We also use a Monte Carlo study to assess the direction of the bias. We observe that the incidental parameter problem in our application is similar to a problem of measurement error in variables, generating attenuation bias in the estimates of the parameters of interest. Primarily, our estimates of γ and ρ are the most affected ones by the incidental parameter problem, and are the ones that benefit the most from the bias correction.

Our model belongs the class of models discussed in Fernandez-Val and Vella (2011). However, we are able to use a simpler bias correction method due to our assumption that shocks are not serially correlated. Furthermore, instead of our iterative procedure it would be possible to estimate our coefficients by full information maximum likelihood (FIML). Although FIML may be more efficient under joint normality, our method requires distributional assumptions weaker than joint normality of u_{it}, v_{it} and α_i; and can include α_i as a fixed, rather than random effects. Due to the large size of the dataset used in this study, we have enough precision to make inference and therefore robustness is our main concern.

Finally, confidence intervals are obtained by bootstrap. Bootstrap is based on 500 replications of our entire estimation procedure. Sampling is done with replacement over i, therefore we use all observed time periods for a given individual.

4.1 Estimates

The estimation results from our recursive algorithm are presented in Tables 3 with log wages as the outcome of interest, and in Table 4 where log income is the outcome of interest. The first three columns of each table present estimates for the sample of men, the next three columns for the sample of women. For both the male and female sample we first present the results for the selection equation, followed by the outcomes equation for Souther Italy, and the outcome equation for Northern Italy. Throughout we use our mobility indicators (average number of moves between employers per year, and an indicator equal to one if the worker has never changed 1-digit industries) as our z variables, excluding them from the outcome equations.[24] The outcomes are functions of ability (α) and the observable time varying characteristics (x) which are: experience and experience squared, tenure and its squared, years in the North and its squared, as well as indicators for occupation (blue collar, white collar and managerial occupation), part time job, year and multi region firm (a firm that has establishments in both North and South). The inclusion of year fixed effects in the selection equation controls for potential migrants' time-varying outside options, in particular also the option to emigrate from Italy.[25]

The central question posed in this paper is whether there is selection on ability, where "ability" refers to the time-invariant characteristics of each worker that contribute toward a worker's wage and income. We find strong evidence that there is negative selection on ability for both men and women. Southern Italians with a lower fixed effect in the wage equation are more likely to migrate to the North. The point estimate for men implies that a one standard deviation increase in ability, as it matters for wages, decreases the annual probability of migrating to Northern Italy by 7 percent (on average the annual probability decreases from 6.7 to 6.2 percent). For women a one standard deviation higher ability results in the annual probability of migrating decreasing by 2.4 percent (from 3.4 to 3.3 percent). Selection on ability is considerably more pronounced when measured in terms of income. A one standard deviation increase in ability, as it matters for income, decreases the annual probability of migrating to Northern Italy by 36 percent (from 6.7 to 4.3 percent) for men and by 30 percent for women (from 3.4 to 2.4 percent).

[24] An individual's propensity to change jobs and her wage may of course be correlated for numerous reasons. Individual heterogeneity likely plays a prominent role in explaining that correlation. For example, individuals with low job attachment may also be less productive (or picky) in their work and therefore receive lower wages (in the spirit of extensions of the McCall,1970, model). Quality of job-worker matches (Jovanovic, 1979) and early career dynamics (Neal, 1999) are also likely important.

Our wage equations include individual fixed effects, controls for tenure and potential experience, thus accounting for most conceivable sources of correlation between our mobility indicators and wages, other than selection.

[25] See Fernandez-Huertas (2013) for a discussion of this approach, and Bertoli and Fernandez-Huertas (2013) for a discussion of multilateral resistance to migration.

Inspection of the estimates for the selection equations further shows that the probability of migrating is decreasing in tenure, increasing in the duration spent in the North, and slightly increasing in potential experience. White collars and in particular managers are more likely than blue collar workers to migrate, part-time less likely. Those employed in a firm that has establishments in both regions are more likely to migrate. Our excluded variables are highly significant in the selection equations. Though the probability of migrating is, somewhat surprisingly, decreasing in the cumulative average number of moves between employers per year. We assume that once we control for the worker fixed effect and for tenure, which of course depends on how often a worker changes employer, these do not directly affect wages or employment.[26]

Turning to the outcome equations it is worth noting that much existing work on migration decisions assumes that there is no selection bias in the observed wages due to migration (see for example, Chiquiar and Hanson, 2005, and Fernandez-Huertas, 2011). In contrast, we find clear evidence of selection bias due to migration in the wages in the South. The coefficients on the selection correction term are consistently negative and statistically significant for both men and women in wages in the South, and women in the income specification. The implication is that the probability of migrating is decreasing if the individual experiences a positive transitory shock to wages or employment in the South. Ignoring selection bias would result in an overestimate of the counterfactual wages (and income) of migrants had they remained in the South, resulting in a bias toward finding positive selection of migrants.

Evidence of selection bias in migrant wages in Northern Italy is more mixed. The

[26]The model described in Section 2 makes explicit the mechanism that generates selection of migrants. Workers decide to migrate according to the present value of their future labor market income. Returns to the workers' characteristics differ across locations and therefore workers with different characteristics take different migration decisions. If the true data generating process were a model as the Roy model described in Section 2, it would not be necessary to include an instrument in the selection equation. Moreover, we report evidence that supports the structural model described in Section 2; in our application, the Roy model does have a good performance fitting the data. However, for the sake of robustness and in the spirit of reduced form estimation, we also include an instrument in the selection equation, which is excluded from the outcome equations. However, we have also estimate our baseline specification without including en external source of variation in the selection equation and results do not change significantly.

coefficient on the selection correction term (the inverse Mills ratio) for the wage equation in the North is negative and significant for men: a positive shock in North increases probability of migrating. However, it is not significant for men in the income equation and for women in the wage equation; and it is positive for women in the income equation. In sum, the evidence suggests that potential migrants respond to transitory shocks in the source region, but it is less clear whether they respond strongly to transitory shocks in the destination region. This is consistent with the idea that migrants may face considerable uncertainty about their economic prospects pre-migration.

Consistent with our finding of negative selection in terms of ability, we find that the returns to ability for migrants are lower in the North than in the south of Italy. ρ is found to be significantly lower than one for male and female workers in the models for wage and income. This finding is remarkably stable across specifications, we reject the null of $\rho = 1$ in all the robustness checks presented in the paper (see Tables 7, 8, 9 and 10 in the Appendix). It is worth highlighting that our unique dataset allows us to track workers before and after the migration decision. This information is generally not available in most of the dataset used in the literature and is fundamental for the identification of a price differential of ability.

Wages and income are increasing and slightly concave in potential experience and increasing and concave in tenure in both North and South, for men and women. The return, however, are remarkably low, though they are significantly higher when individual fixed effects are not included. Despite controlling for individual fixed effects, in both the South and North blue collar workers make lower wages and a lot lower income than white collar workers, who make a lot less than managers. The fact that income differentials are greater than wage differentials is a result of higher paid professions also providing more stable employment. Part-time workers are paid higher wages, though they are of course employed less (full–time equivalent) weeks per year. There is a wage and income premium for working in firms with establishments in multiple regions.

Wages and income for male migrants in Northern Italy increase (at a decreasing rate) with time spent in the North, by 4 and 22 percent respectively the first year, evidence of the importance of assimilation for migrant outcomes. There is no statistically significant effect for female migrants though. Interestingly, time spent in the North also has a positive effect on wages in the South for return migrants, but only for long migration spells. In the case of males, only migrants who spent at least eight years in the North have wages in the South which are higher than the wage that they would have received if they had not migrate.

We conduct numerous robustness checks reported in the Appendix, none of which change our results qualitatively. In Tables 7 and 8, we present results using a smaller sample that exclude return migrants from the analysis. We recalculate our coefficients for male and female workers using the model of wage as well as the model of income. In Tables 9 and 10 we also report results with a different definition of the North and the south of Italy that excludes the central Italian provinces (Lazio, Marche and Umbria) to avoid identifying commuters as migrants, see Figure 2. We report results for male and female workers, including and excluding return migrants for both models (wage and income).

4.2 Selection of Migrants

The estimated actual (for non-migrants) and counterfactual (for migrants had they stayed in the South) weekly wage and income densities for Southern Italian workers are in Figure 4. We only show the distributions for men since, as our discussion of the estimates suggests, the results for women are similar though less pronounced. For men we find considerable differences between the counterfactual wage and income densities for migrants and the actual densities for non-migrants. Male migrants are disproportionately drawn from the lower half of the wage distribution in Southern Italy, providing evidence of negative selection while there is intermediate selection of migrant if we take employment opportunities into account.

The densities of our estimates of the contribution of observable time-varying characteristics to wages and income ($\beta_s x$) are shown in Figure 5. They suggest that there is intermediate selection of migrants on observable characteristics in terms of income, and slightly negative selection of migrants when measured in terms of wage. Figure 6 presents our estimates of the distribution of ability, from wage and income equations, for workers in the South. Male migrants are disproportionately drawn from the lower half of the ability distribution. The degree of negative selection is more pronounced in terms of income, highlighting both the importance of ability and employment opportunities for characterizing the selection of migrants.

For comparison, we present the actual and counterfactual wage distributions for non-migrants and migrants assuming that selection operates only through observable characteristics, Figure 8 using the methodology of Chiquiar and Hanson (2005), and selection on unobservables, Figure 9 using the methodology of Fernandez-Huertas (2011). Negative selection of migrants is evident using either methodology. However, this negative selection is least pronounced when allowing for selection only on observables, and it is more pronounced when accounting for differences in unobservables. In our application both methodologies underestimate the degree of negative selection of migrants from the wage distribution by failing to account for the fact that time invariant unobserved characteristics may enter the selection equation.

Caution is required in interpreting our results since, unlike in a typical Mincer equation, we do not observe an individual's education attainment. However, we have precise information on occupation, age and experience which capture most of the variation of the wage explained by observables. In Table 11 we present a variance decomposition of wages using a sample of Italian employees in 2007 from the EU-SILC data on Italy. We find that less than 1.5 percent of the variation of residuals from a wage regression with a specification as close as possible to the one used in the paper (which controls for experience, occupation and further individual and firm characteristics) is explained by education. In Italy educational attainment captures only a small fraction of what constitutes an individual's ability, as well as explaining only a small fraction of overall wage dispersion.

Around half of those who migrate to the north of Italy return within our sample. The question is whether, as suggested by Borjas and Bratsberg (1996), return migration reinforces the original selection of migrants, or not. Table 5 shows mean and median outcomes (wages and income), ability and returns to observables - based on the outcome equations in the South - for non-migrants, migrants who do not return to South and return migrants. Male migrants are negatively selected from the ability distribution, for both wages and income, and return migration reinforces that selection: migrants who do not return to the South are on average of much lower ability than those who return. For observable characteristics that affect wages there is no such clear pattern. In terms of income, however, migrants are positively selected; a pattern which is reinforced by return migration. The pattern of selection for wages and income is the result of negative selection on ability and positive (or intermediate) selection on observed characteristics.

4.3 Returns to Migration

Figures 7a and 7b show predicted (*ex ante*) returns to migration for male migrants and non-migrants by duration of the migration experience. As predicted by the Roy model expected returns for migrants are consistently higher than those for non-migrants. The expected returns of migration in terms of wages in the first year are 1.5 percent for migrants and -3 percent for non-migrants and then both grow monotonically (at a decreasing rate) with duration in the North. For income the first year's returns are negative, they are close to zero for migrants in the second year and -13 percent for non-migrants. Thereafter they continue to grow monotonically (at a decreasing rate).

Figures 7c and 7d show estimated actual (*ex post*) returns to migration for migrants (actual wages in the North minus counterfactual wages in the South). Annual returns to migration are always positive in terms of wages: 5, 15 and 21 percent in the first, fifth and tenth year respectively. In terms of income they are -7, 8, 33, and 50 percent in the first, second, fifth and tenth year respectively. Around half the gains from migration (after the first year) are due to higher wages, and the other half due to better labor market attachment. The fact that the income gains due to migration are negative in the first year in part reflects that most migration experiences involve an interruption in employment.

Two key hypotheses about return migration is that they (1) reflect uncertainty about returns to migration, and (2) are part of a human capital acquisition strategy. The positive returns to a migration experience for return migrants in Southern Italy support the second hypothesis. The results presented in Table 5 provide support for the first hypothesis. We define a return migrant as someone currently in the North, but return to South within sample. Returns to migration are significantly higher for those who never return. In terms of wages in the first year the returns are 8 percent for non-returnees and 5 percent for returnees, in the fifth year 17 and 10 percent. The gains in income are 7 percent for non–returnees and -13 percent for returnees in the first year; and 52 percent for non-returnees and 4 percent for returnees after five years. The evidence is clearly consistent with the idea that a lot of return migration is the result of a disappointing migration experience, and thus we should not be surprised if it is not necessarily associated with wage or income gains.

5 Conclusions

Understanding migration patterns, who migrates and why they do so, is critical for understanding the impact on source and destination regions and countries, as well as informing the feasibility of policy to affect these decisions. In this paper we use the fact that we have multiple observations on migrants, from poor Southern Italy to wealthy Northern Italy, to identify the importance of ability for selection of migrants and returns to migration. We propose and implement a novel iterative estimation method for a switching regression model with the same worker-specific source of unobserved heterogeneity (worker fixed effects) present in the selection and both outcome equations.

We find that differential returns to unobserved worker characteristics ("ability") and differences in employment opportunities between regions are important determinants of migration decisions. We estimate that the returns to ability are lower in the North than in the South and accordingly migrants tend to be drawn from the lower-end of the ability distribution, even more so if we also account for changes in employment. Differential returns to observable characteristics are far less important, which may explain why studies of migration decisions who focus of these have, despite its obvious intuitive appeal, not found strong support for the predictions of the Roy model. Both assimilation and selection are important as the returns to migration rise with duration of the migration experience. Return migration is an important phenomenon in Italy and reinforces the original negative selection of migrants. This is consistent with the idea that migrants face considerable uncertainty about their income in the north of Italy, resulting in a lot of marginal migrants who return as their expectations are disappointed. Return migrants, who spent a significant amount of time in the North, enjoy positive returns to this migration experience on their return to the South, suggesting a role for migration as a human capital acquisition strategy.

The focus of this paper has been on individual migration decisions, ignoring general equilibrium effects. A clear direction for subsequent research is to examine the factors that affect the volume of migration (and return migration) flows and the associated consequences. For example, how migration flows are affected by the business cycles in source and destination location, and how, in turn this affects relative wages and employment in both locations.

References

[1] Abramitzky, Ran (2008). "The Limits of Equality: Insights from the Israeli Kibbutz," Quarterly Journal of Economics, MIT Press, vol. 123(3), pages 1111-59.

[2] Ran Abramitzky, Leah Platt Boustan and Katherine Eriksson (2012). "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration," American Economic Review, vol. 102(5), pages 1832-56, August.

[3] Ambrosini, J. William and Giovanni Peri (2012). The determinants and the selection of Mexico-U.S. migrants." World Economy, 111–151.

[4] Arellano, Manuel and Olympia Bover (2002). "Learning about migration decisions from the migrants: Using complementary datasets to model intra-regional migrations in Spain," Journal of Population Economics, Springer, vol. 15(2), pages 357-80.

[5] Barrett, Alan and Jean Goggin (2010). "Returning to the Question of a Wage Premium for Returning Migrants," National Institute Economic Review, National Institute of Economic and Social Research, vol. 213(1), pages R43-51.

[6] Bertoli, Simone and Jesús Fernández-Huertas Moraga (2013). "Multilateral Resistance to Migration," Journal of Development Economics, Elsevier, vol. 102, pages 79-100.

[7] Bertoli, Simone, Jesús Fernández-Huertas Moraga and Francesc Ortega (2013). "Crossing the Border: Self-Selection, Earnings and Individual Migration Decisions," Journal of Development Economics, Elsevier, vol. 101, pages 75-91.

[8] Bonifazi, C., F. Heins, S. Strozza and M. Vitiello (2009), "The Italian transition from emigration to immigration country", IDEA Working Papers No. 5, March.

[9] Borjas, George J (1987). "Self-Selection and the Earnings of Immigrants," American Economic Review, American Economic Association, vol. 77(4), pages 531-53.

[10] Borjas, George (1991). "Immigration and Self-Selection," in John Abowd and Richard Freeman, e.d., *Immigration, Trade and the Labor Market*, NBER, pp. 29-76.

[11] Borjas, George (2008). "Labor Outflows and Labor Inflows in Puerto Rico." Journal of Human Capital, 2(1), pages 32-68.

[12] Borjas, George and Bernt Bratsberg (1996). "Who Leaves? The Outmigration of the Foreign-Born," Review of Economics and Statistics, vol. 78(1), pages 165-76.

[13] Borjas, George, Stephen Bronars, and Stephen Trejo (1992). "Self-Selection and Internal Migration in the United States," Journal of Urban Economics, Elsevier, vol. 32(2), pages 159-85.

[14] Borger, Scott (2010). "Self-Selection and Liquidity Constraints in Dfferent Migration Cost Regimes," Working Paper, Office of Immigration Statistics.

[15] Caponi, Vincenzo (2011). "Intergenerational transmission of abilities and self- selection of Mexican immigrants." International Economic Review, 58: 523—547.

[16] Carrington, William and Enrica Detragiache (1998) "How Big is the Brain Drain?," IMF Working Papers 98/102, International Monetary Fund.

[17] Chiquiar, Daniel and Gordon H. Hanson (2005). "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," Journal of Political Economy, University of Chicago Press, vol. 113(2), pages 239-81.

[18] Co, Catherine, Ira Gang, and Myeong-Su Yun (2000). "Returns to Returning: Who Went Abroad and What Does it Matter?" Journal of Population Economics, vol. 13(1), pages 57-79.

[19] Contini B., Filippi M., Malpede C. (2000), "Safari tra la giungla dei salari. Nel Mezzogiorno si lavora meno?", Lavoro e Relazioni Industriali, vol.2.

[20] de Coulon, Augustin and Matloob Piracha (2005). "Self-selection and the performance of return migrants: the source country perspective," Journal of Population Economics, vol. 18, pages 779-807.

[21] Dahl, G. B. (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," Econometrica, Econometric Society, vol. 70(6), pages 2367-420.

[22] Del Boca, Daniela and Alessandra Venturini (2005). "Italian Migration," in K. F. Zimmermann, e.d., *European Migration - What Do We Know?*, Oxford University Press.

[23] Docquier, Frédéric and Hillel Rapoport (2012). "Globalization, Brain Drain and Development," Journal of Economic Literature, vol. 50(3), pages 681-730.

[24] Dustmann, Christian and Yoram Weiss (2007). "Return Migration: Theory and Empirical Evidence from the UK," British Journal of Industrial Relations, London School of Economics, vol. 45(2), pages 236-56.

[25] Dustmann, Christian, Itzhak Fadlon, and Yoram Weiss (2011). "Return Migration, Human Capital Accumulation, and the Brain Drain," Journal of Development Economics, 95(1), 58-67.

[26] Fernández-Huertas Moraga, Jesús (2011). "New Evidence on Emigrant Selection," Review of Economics and Statistics, 93:72—96.

[27] Fernández-Huertas Moraga, Jesús (2013). "Understanding different migrant selection patterns in rural and urban Mexico," Journal of Development Economics, 103, pp.182–201.

[28] Fernandez-Val, Ivan and Frank Vella (2011) "Bias Corrections for Two-Step Fixed Effects Panel Data Estimators," Journal of Econometrics 163(2), pp. 144-162.

[29] Gilmour, David (2011). "The Pursuit of Italy: A History of a Land, Its Regions, and Their Peoples," Farrar, Straus and Giroux.

[30] Goria, A. and A. Ichino (1994), " Flussi Migratori e Convergenza tra Regioni Italiane", Lavoro e Relazioni Industriali, 3, luglio-settembre, 3-50.

[31] Grogger, Jeffrey and Gordon Hanson (2011). "Income maximization and the selection and sorting of international migrants," Journal of Development Economics, Elsevier, vol. 95(1), pages 42-57.

[32] Hahn, Jinyong and Whitney Newey (2004). "Jackknife and Analytical Bias Reduction for Nonlinear Panel Models," Econometrica, Econometric Society, vol. 72(4), pages 1295-319.

[33] Hanson, Gordon and Craig McIntosh (2010). "The Great Mexican Emigration," Review of Economics and Statistics, vol. 92(4), pages 798-810.

[34] Heckman, James J. (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," Annals of Economic and Social Measurement, vol. 5, pages 475-92.

[35] Heckman, James J. (1979). "Sample Selection Bias as a Specification Error," Econometrica, Econometric Society, vol. 47, pages 153-61.

[36] Heckman, James J and Bo E. Honore (1990). "The Empirical Content of the Roy Model," Econometrica, Econometric Society, vol. 58(5), pages 1121-49.

[37] Hoderlein, Stefan and Halbert White (2009). "Nonparametric Identification in Nonseparable Panel Data Models with Generalized Fixed Effects," Boston College Working Papers in Economics 746, Boston College Department of Economics.

[38] Ibarraran, Pablo and Darren Lubotsky (2007). "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census," in George Borjas, e.d., Mexican Immigration to the United States, NBER, pp. 13-56.

[39] Jovanovic, Boyan (1979). "Job Matching and the Theory of Turnover." Journal of Political Economy, Vol. 87(5), pp. 972–990.

[40] Kaestner, Robert and Ofer Malamud (2013). "Self-selection and international migration: New evidence from Mexico." Review of Economics and Statistics, forthcoming.

[41] Kennan, John and James Walker (2011). "The Effect of Expected Income on Individual Migration Decisions," Econometrica, Econometric Society, vol. 79(1), pages 211-51.

[42] Lacuesta, Aitor (2006). "Emigration and human capital: who leaves, who comes back and what difference does it make?," Banco De Espana.

[43] Lee, Lung-Fei (1976). "Two-Stage Estimations of Limited Dependent Variable Models," Ph.D. Thesis, University of Rochester, Rochester, N.Y.

[44] Lemieux, Thomas (2006). "Postsecondary Education and Increasing Wage Inequality," American Economic Review, American Economic Association, vol. 96(2), pages 195-99.

[45] Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press.

[46] Massey, Douglas S., Jorge Durand, and Nolan J. Malone (2003). "Beyond Smoke and Mirrors: Mexican Immigration in an Era of Economic Integration," Russell Sage Foundation Publications.

[47] Mattoo, Aaditya, Ileana Neagu, and Caglar Ozden (2008). "Brain waste? Educated immigrants in the US labor market," Journal of Development Economics, Elsevier, vol. 87(2), pages 255-69.

[48] McCall, John (1970). "Economics of Information and Job Search." Quarterly Journal of Economics, Vol. 84(1), pp. 113–126.

[49] McKenzie, David, John Gibson, and Steven Stillman (2010). "How Important is Selection? Experimental Versus Non-Experimental Measures of the Income Gains From Migration." Journal of the European Economic Association, European Economic Association, vol. 8(4), pages 913-45.

[50] McKenzie, David and Hillel Rapoport (2010). "Self-Selection Patterns in Mexico-US Migration: the Role of Migration Networks", The Review of Economics and Statistics, vol. 92(4), pages 811-21.

[51] Meghir, Costas. and Pistaferri, Luigi. (2004). "Income Variance Dynamics and Heterogeneity," Econometrica, 72(1), pp. 1-32.

[52] Neal, Derek (1999). "The Complexity of Job Mobility among Young Men." Journal of Labor Economics, Vol. 17(2), pp. 237–261.

[53] Newman, J. and Elizabeth Scott (1948). "Consistent Estimates Based on Partially Consistent Observations," Econometrica, Econometric Society, vol. 16(1), pages 1-32.

[54] Reinhold, Steffen and Kevin Thom (2011). "Migration Experience and Earnings in the Mexican Labor Market," Working Paper.

[55] Roy, A. D. (1951). "Some Thoughts on the Distribution of Earnings," Oxford Economic Papers, New Series, vol. 3(2), pages 135-46.

[56] Sjaastad, Larry (1962). "The Costs and Returns of Human Migration," The Journal of Political Economy, vol. 70(5), pages 80-93.

[57] Thom, Kevin (2010). "Repeated Circular Migration: Theory and Evidence from Undocumented Migrants," Working Paper.

[58] Tunali, Insan (2000). "Rationality of Migration," International Economic Review, Vol. 41, Issue 4, November.

[59] Verbeek, Marno and Theo Nijman (1992). "Testing for Selectivity Bias in Panel Data Models." International Economic Review, vol. 33(3), pages 681-703.

[60] Wooldridge, Jeffrey (1995). "Selection Corrections for Panel Data Models Under Conditional Mean Independence Assumptions," Journal of Econometrics, vol. 68, 115-32.

[61] Zabel, Jeffrey (1992). "Estimating Fixed and Random Effects Models with Selectivity," Economics Letters 40, 269-72.

[62] Zamagni, Vera (1998). *The Economic History of Italy 1860-1990*, Oxford University Press.

Appendix

A Return Migration

A considerable number of migrants return to their source country. We show that it is possible to analyze their return migration decision in the same framework as the original migration decision. Consider an individual i who has already migrated to the destination country and now every period t has the choice whether to stay $R_{it} = 1$ in the destination country $j = n$ or return migrate $R_{it} = 0$ to the original source country $j = s$. As before we assume she makes that decision based on the difference in outcomes y_{ijt} in each country and a one-time return migration cost r_{it}. The return migration decision is carried out according to

$$R_{it} = \begin{cases} 1 & \text{"stay" iff } y_{itn} - y_{its} - r_{it} \geq 0 \\ 0 & \text{"return migrate" iff } y_{itn} - y_{its} - r_{it} < 0 \end{cases}$$

As before we will allow return migration costs r_{it} to be a function of the time-varying x and time invariant α individual characteristics that affect outcomes, as well as other individual characteristics z that are excluded from the outcome equations and unobserved individual-specific time-varying factors, which are assumed to enter additively separable. Then we can rewrite the selection equation for potential return migrants as

$$R_{it} = 1\left(r\left(\alpha_{i,}x_{it}, z_{it}\right) + \omega_{it} \geq 0\right), \tag{10}$$

where ω_{it} is distributed independently of α, x , and z with mean zero and variance σ_{ω}^{2}.

To incorporate the possibility of return migration into our basic model of migration we make the crucial assumption that the same potentially time-varying unobserved factors affect the return migration and migration decisions such that $\omega_{it} \equiv v_{it}$. To provide a better idea for the intuition behind this restriction note that

$$\begin{aligned} v_{it} &= u_{itn} - u_{its} - \zeta_{it}^{M}, \\ \omega_{it} &= u_{itn} - u_{its} - \zeta_{it}^{R}, \end{aligned}$$

where ζ_{it}^{M} are idiosyncratic factors that affect the decision to migrate from source to destination country, and ζ_{it}^{R} are idiosyncratic factors that affect the decision to stay in the destination country and not return migrate. We assume that $\zeta_{it}^{M} \equiv \zeta_{it}^{R}$, implying that these factors are not moving costs *per se*, but rather affect the preference for living in a certain country. This, to us, does not seem like an unduly restrictive assumption.

We can then combine selection equations (3) and (10) into a selection equation describing whether the individual chooses to work in the destination country $N_{it} = 1$ or the source country $N_{it} = 0$

$$N_{it} = 1 \left(m \left(\alpha_{i}, x_{it}, z_{it} \right) D_{i,t-1} + r \left(\alpha_{i}, x_{it}, z_{it} \right) \left(1 - D_{i,t-1} \right) + v_{it} \geq 0 \right), \tag{11}$$

where $D_{it} = 1 \left(y_{it}^{*} = y_{its} \right)$, i.e. if the individual was observed working in the source country last period.

Then the outcome equations conditional on actually being observed contain two control functions, one for migrants and one for return migrants

$$\begin{aligned} y_{itn} \mid_{N=1} &= \rho \alpha_{i} + \beta_{n} x_{it} + \frac{\sigma_{nv}}{\sigma_{v}} \left(D_{it} \frac{\phi \left(m \left(\alpha_{i}, x_{it}, z_{it} \right) \right)}{\Phi \left(m \left(\alpha_{i}, x_{it}, z_{it} \right) \right)} + \left(1 - D_{it} \right) \frac{\phi \left(r \left(\alpha_{i}, x_{it}, z_{it} \right) \right)}{\Phi \left(r \left(\alpha_{i}, x_{it}, z_{it} \right) \right)} \right) + \varepsilon_{itn}, \\ y_{its} \mid_{N=0} &= \alpha_{i} + \beta_{s} x_{it} - \frac{\sigma_{sv}}{\sigma_{v}} \left(D_{it} \frac{\phi \left(m \left(\alpha_{i}, x_{it}, z_{it} \right) \right)}{1 - \Phi \left(m \left(\alpha_{i}, x_{it}, z_{it} \right) \right)} + \left(1 - D_{it} \right) \frac{\phi \left(r \left(\alpha_{i}, x_{it}, z_{it} \right) \right)}{1 - \Phi \left(r \left(\alpha_{i}, x_{it}, z_{it} \right) \right)} \right) + \varepsilon_{its}, \end{aligned}$$

where the coefficient on the control functions for potential migrants and potential return migrants are the same. Note that the set of observables can include variables that depend on whether you have been a migrant or not, for example, years spent in destination country. Hence, (return) migration can affect outcomes on account of differences in factor prices between the source and destination country, the returns to sorting, and because of differences in observables that are a function of the (return) migration decision.

B Monte Carlo Studies

In our Monte Carlo study we present a simplified version of the model used in the empirical section. The model design is:

$$
\begin{aligned}
y_{it} &= \begin{cases} x_{it}\beta_S + \alpha_i + u_{Sit}, \ if \ y^* > 0 \\ x_{it}\beta_N + \rho\alpha_i + u_{Nit}, \ if \ \ y^* \leq 0 \end{cases} \\
y_{it}^* &= x_{it}\theta + \eta z_{it} + \alpha_i\gamma + \varepsilon_{it} \\
\begin{pmatrix} u_{\mathrm{Si}\,t} \\ u_{Nit} \\ \varepsilon_{it} \end{pmatrix} &\sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_S & 0 & \sigma_{S\varepsilon} \\ 0 & \sigma_N & \sigma_{N\varepsilon} \\ \sigma_{S\varepsilon} & \sigma_{N\varepsilon} & \sigma_\varepsilon \end{pmatrix} \right) \\
\alpha_i &\sim N(0, \sigma_\alpha), \ cov(x_{it}, \alpha_i) \neq 0 \\
N &= 1000, \ T = (5, 10, 15, 20) \\
\gamma_0 &= 1, \ \beta_S = 2, \ \beta_N = 2, \ \theta = 3, \ \eta = 5, \ \rho = 0.5
\end{aligned}
$$

We present results for estimates in 100 samples of 1000 individuals each. Two different types of estimators are reported, the values resulted from the iterative algorithm described in section 4 and their bias corrected versions. We describe the evolution of the estimators when T grows, reporting results for $T = 5, 10, 15$ and 20. Table 6 gives the Monte Carlo results for the estimators of γ, β_S, β_N, ρ, θ and η.

In Figure 10, we the performance of our estimation strategy recovering the distribution of the size of the effect of the unobserved fixed heterogeneity in the selection equation $(ie : \gamma)$. We find a significant difference in performance between the non bias corrected and the bias corrected estimators. The distribution of the bias corrected estimate of γ is centered around the true value, while the non corrected one is not. There is no significant improvement between $T = 15$ and $T = 20$, which is reassuring given the time dimension of the panel used in this paper.

In Figure 11, we present the performance of our estimation strategy recovering the distribution of the size of the effect of the unobserved fixed heterogeneity in the outcome equation in the North $(ie : \rho)$. As before, we find a significant difference in performance between the non bias corrected and the bias corrected estimators. The distribution of the bias corrected estimate of γ is centered around the true value, while the non corrected one is not. Once again, there is no significant improvement between $T = 15$ and $T = 20$.

In Figure 12 we show kernel densities of $\hat{\theta}$, $\hat{\eta}$, $\hat{\beta}_S$ and $\hat{\beta}_N$, estimated with the iterative method in 100 samples of 1000 individuals with $T = 15$. We report the densities of the non bias corrected estimates and the bias corrected estimates. In Figure 12 we observe that all coefficients converge in distribution correctly. The bias correction results in significant improvements in the estimates of β_N, θ and η. It does not generate differences in the estimate of β_S.

Figure 1: Differences between the North and the South of Italy

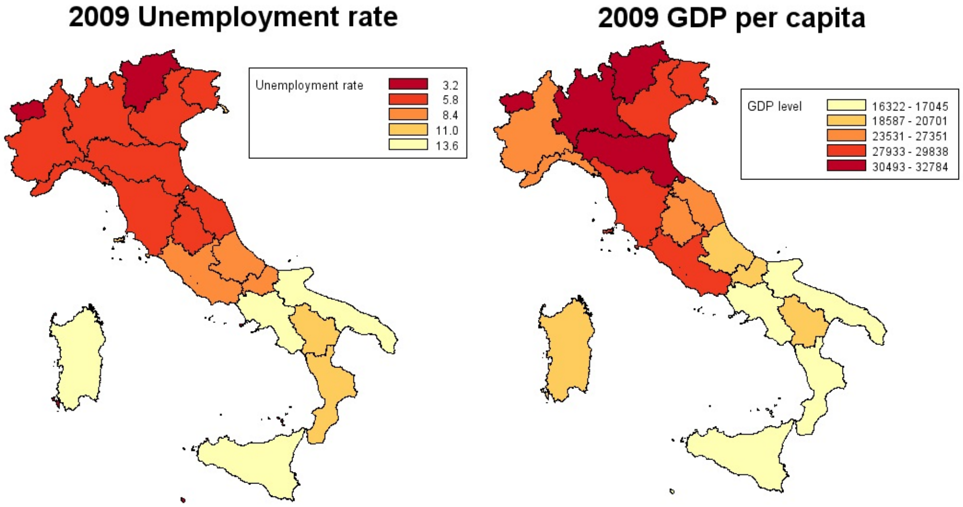


Figure 2: Definition of North and South of Italy

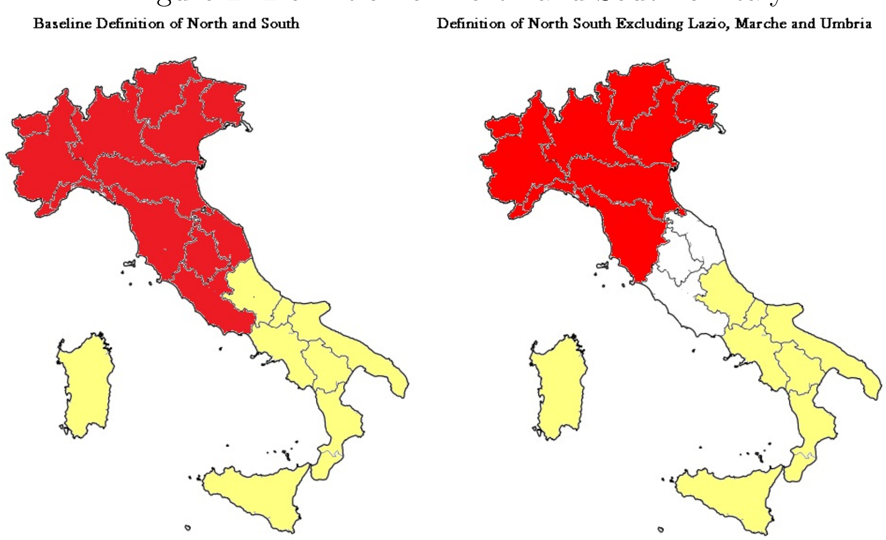


Figure 3: Comparison Between WHIP and IT-SILC

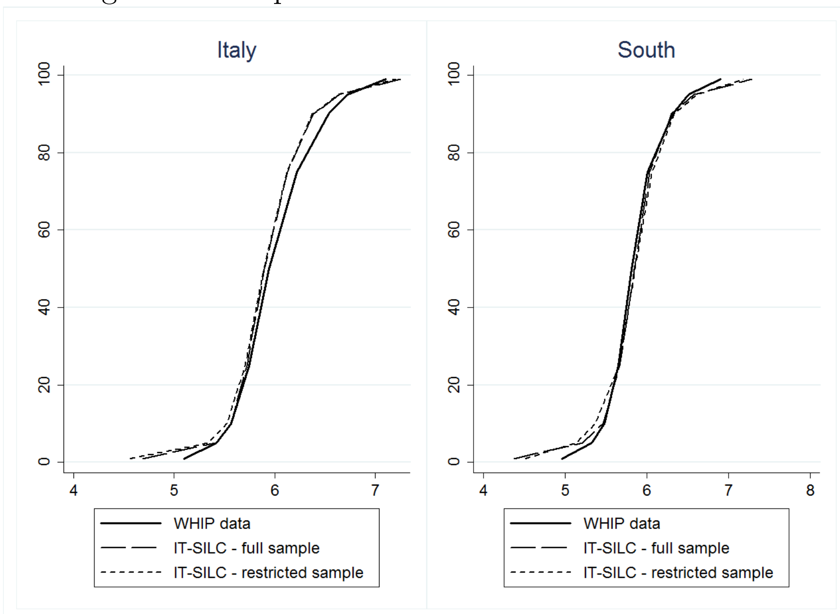
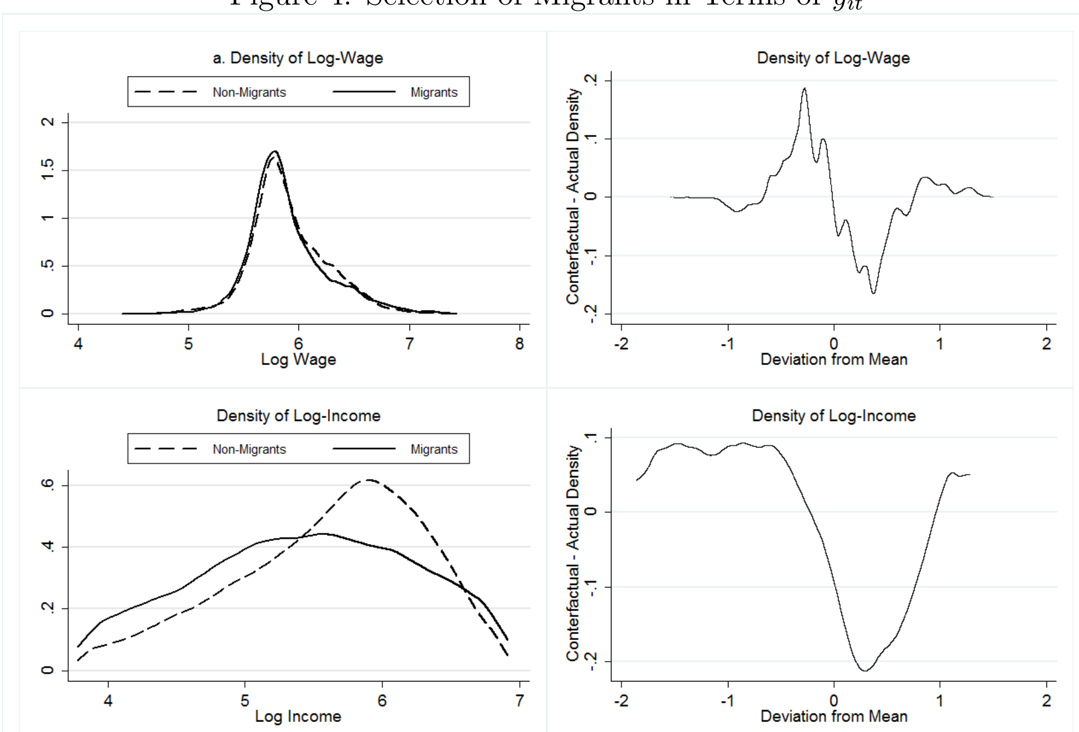
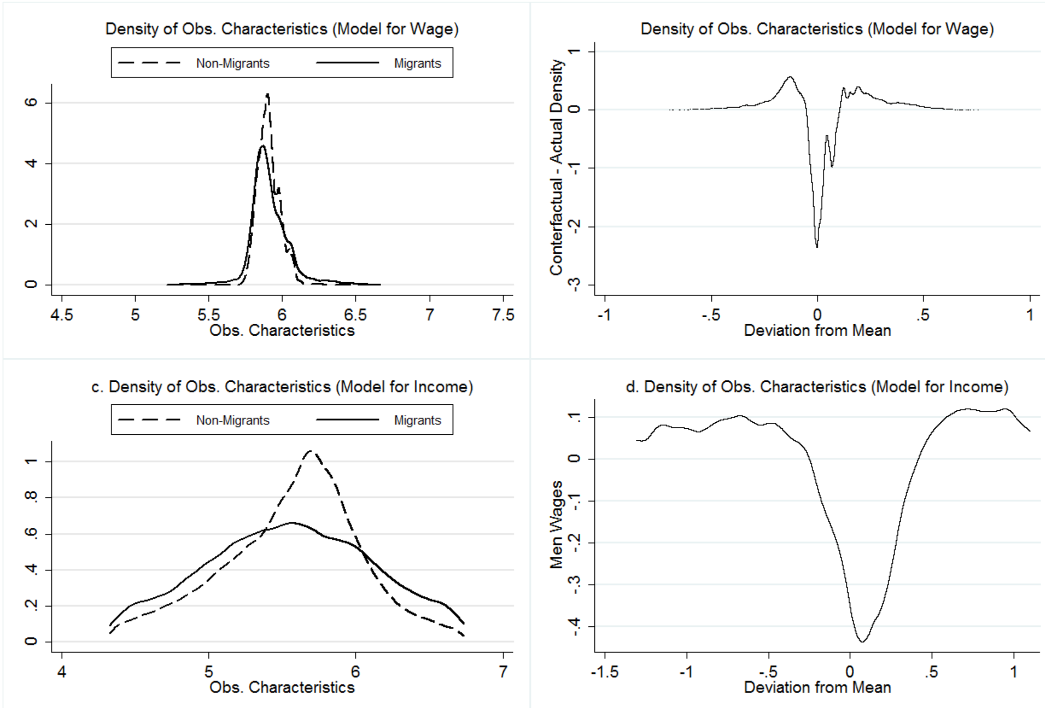


Figure 4: Selection of Migrants in Terms of y_{it}



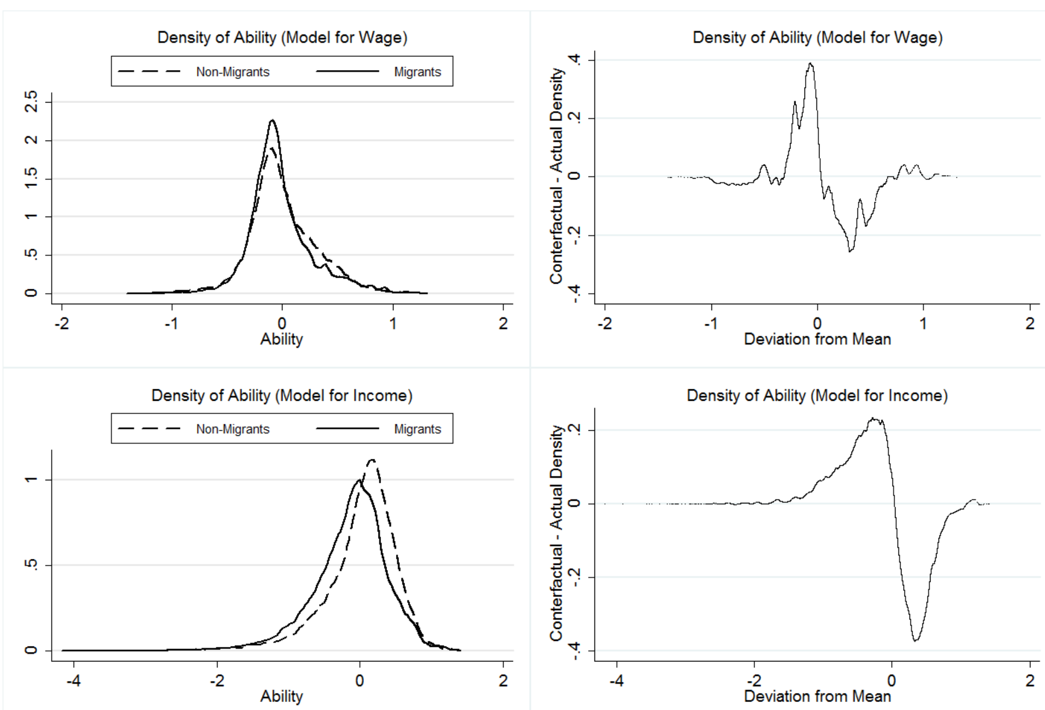
Figures a and b present the distribution of log-wage and log-income for migrants and stayers (only males).
Figures d and c present differences in densities between migrants and stayers (only males).

Figure 5: Selection of Migrants in Terms of βx_{it}



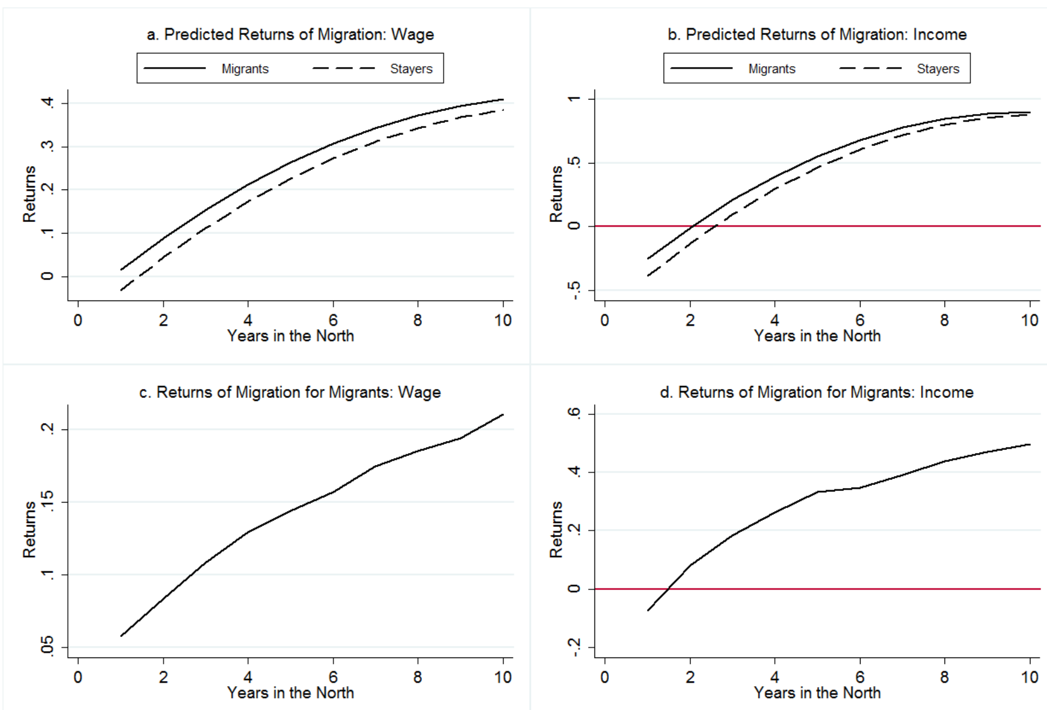
Figures a and b present the distribution of obs. characteristics for migrants and stayers (only males).
Figures d and c present differences in densities between migrants and stayers (only males).

Figure 6: Selection of Migrants in Terms of α_i



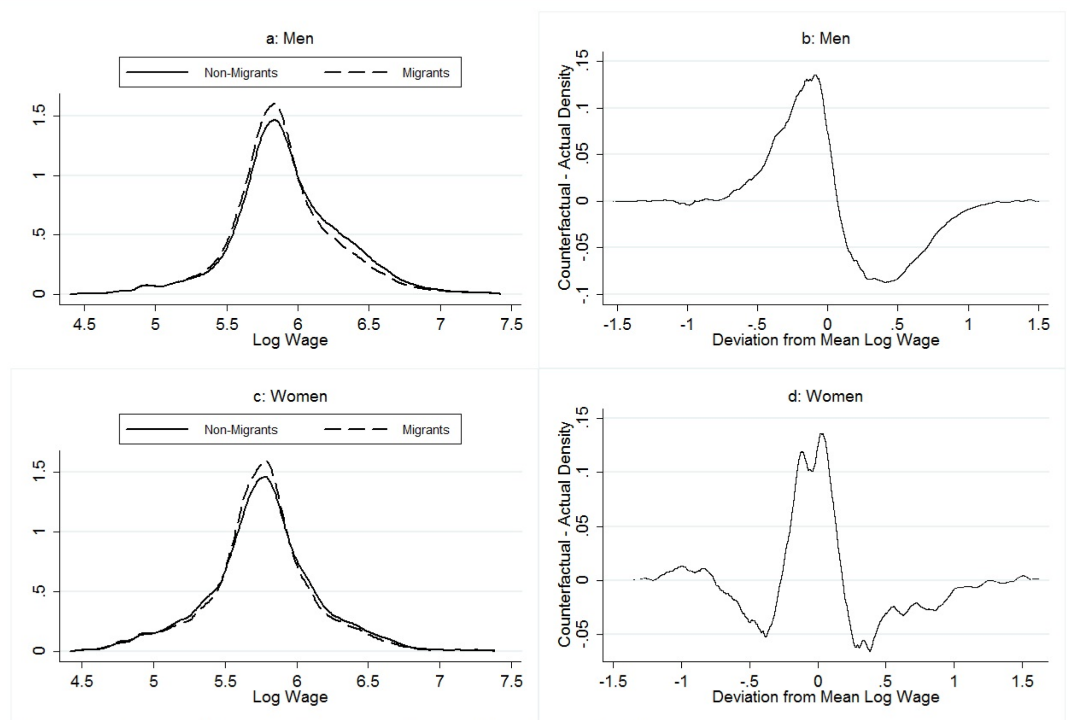
Figures a and b present the distribution of ability for migrants and stayers (only males).
Figures d and c present differences in densities between migrants and stayers (only males).

Figure 7: Gains from Migration for Males



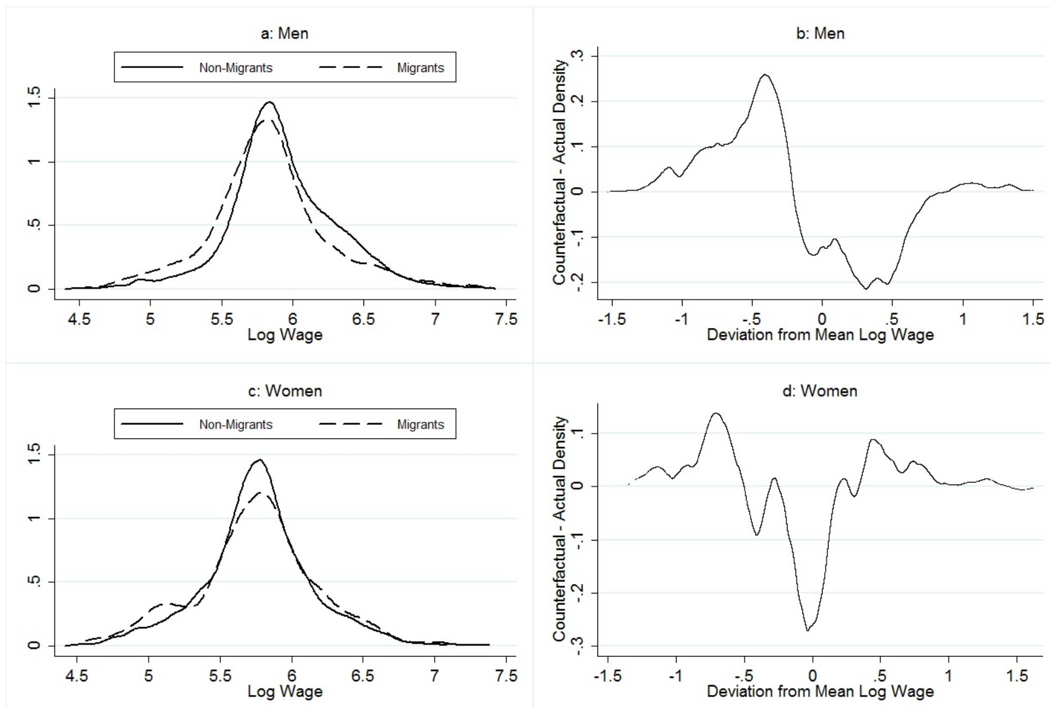
Figures a and b present wage-predictions before migration (predicted wage in the North - predicted wage in the South)
Figures c and d present estimates of returns (realized wage in the North - counterfactual wage in the South)

Figure 8: Selection on Observable Characteristics (Chiquiar and Hanson, 2005)



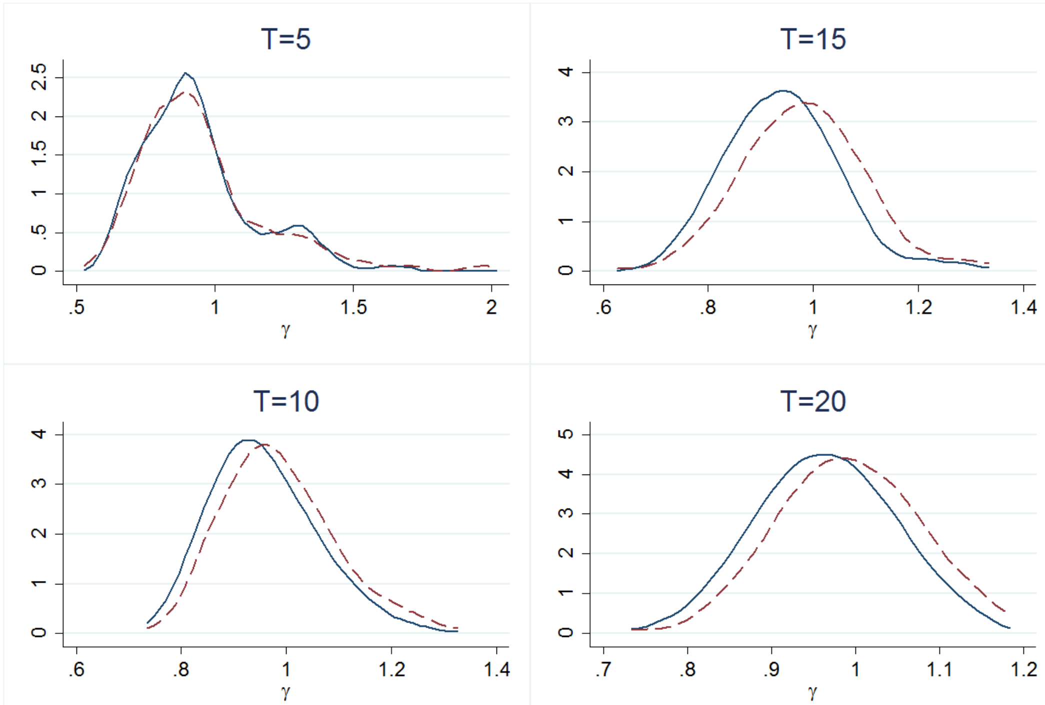
Figures a and c.: actual (non-migrant) and counterfactual (migrant) wage densities.
Figures b and d.: migrant minus non-migrant wage density.

Figure 9: Selection on Unobservable Characteristics (Fernandez-Huertas, 2011)



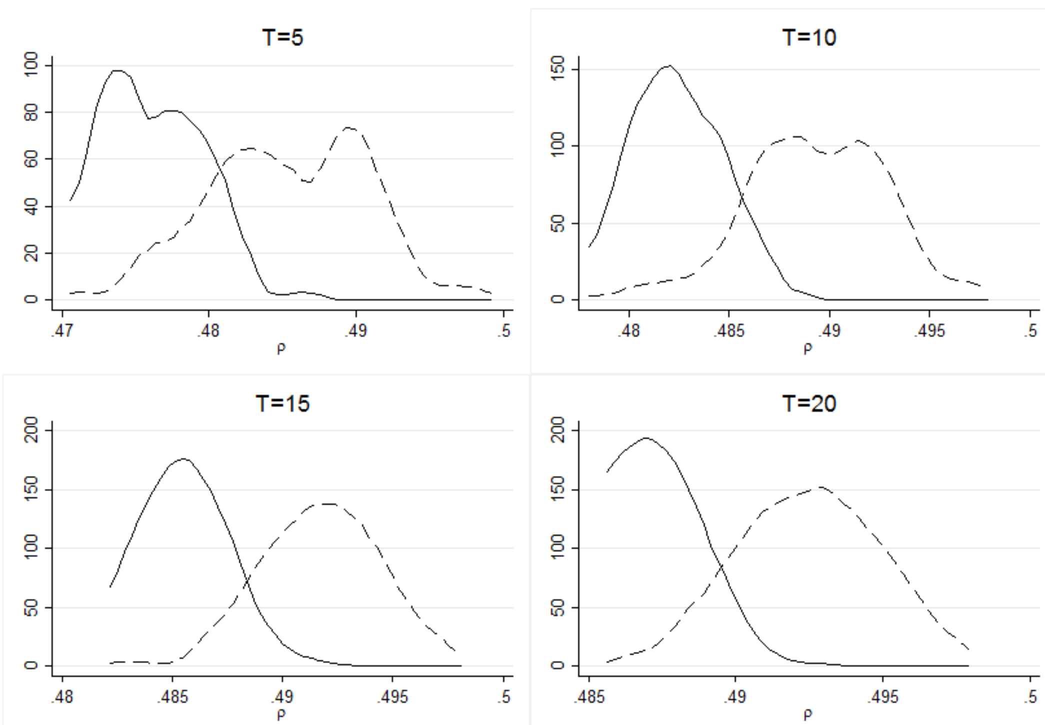
Figures a and c.: actual (non-migrant) and counterfactual (migrant) wage densities.
Figures b and d.: migrant minus non-migrant wage density.

Figure 10: Monte Carlo Simulations: Kernel Density of γ



kernel = epanechnikov, bandwidth = 0.05. Discontinuos line corresponds to the Bias Corrected γ.

Figure 11: Monte Carlo Simulations: Kernel Density of ρ



kernel = epanechnikov, bandwidth = 0.05. Solid line is the non-corrected coefficients and dashed is the bias corrected ones.

Figure 12: Monte Carlo Simulations: Densities of θ, η, β and $cov(u, \epsilon)$

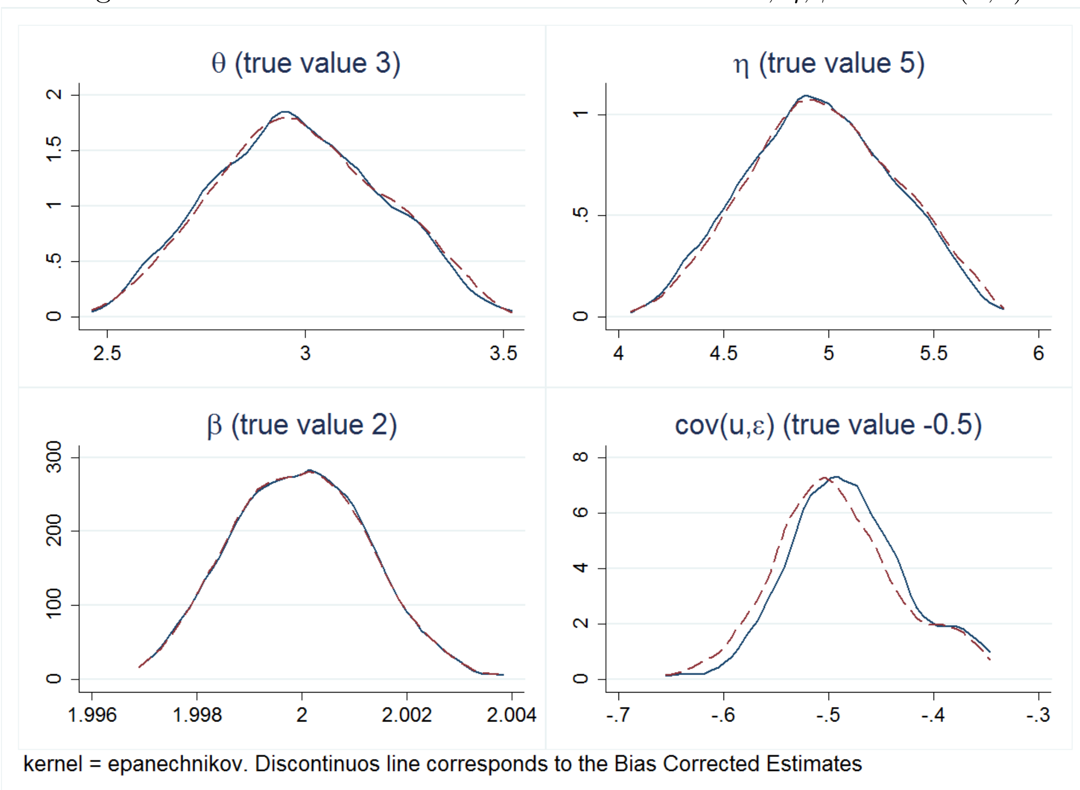


Table 1: Summary Characteristics of Spread of Distributions in log-wages and log-income in the South and the North of Italy

Log-Wage	Place of Work		
	South	North Including Migrants	North Only Natives
	Male Workers		
Mean	5.920	6.106	6.138
Variance	0.150	0.141	0.135
Gini coefficient	0.37	0.034	0.034
Variance not Explained by Obs. Characteristics	0.102	0.084	0.082
	Female Workers		
Mean	5.758	5.913	5.924
Variance	0.150	0.126	0.123
Gini coefficient	0.037	0.032	0.032
Variance not Explained by Obs. Characteristics	0.116	0.088	0.086
Log-Income	Place of Work		
	South	North Including Migrants	North Only Natives
	Male Workers		
Mean	5.566	5.932	6.006
Variance	0.784	0.528	0.432
Gini coefficient	0.081	0.059	0.052
Variance not Explained by Obs. Characteristics	0.526	0.347	0.289
	Male Workers		
Mean	5.238	5.556	5.595
Variance	0.855	0.664	0.610
Gini coefficient	0.093	0.074	0.070
Variance not Explained by Obs. Characteristics	0.556	0.441	0.408

Note: Statistics are based on all workers in Italy (south and north). Wages and income are defined as the weekly average in a given year, they are in euros and are normalized to the 2004 sample mean.

Table 2: Descriptive Statistics on the estimation sample

	Male		Female	
	South	North	South	North
Potential Experience	15.96	15.60	13.87	14.09
Tenure (years)	3.54	2.17	3.29	2.19
Duration in North (years)		2.83		2.59
Blue collar workers (%)	79.28	84.37	55.53	62.68
White collar wokers (%)	20.57	15.27	44.37	36.92
Managers (%)	0.15	0.36	0.10	0.40
Part-time workers (%)	2.49	1.98	16.95	27.05
Multi-Region Firms (%)	11.72	21.47	8.40	23.69
Average Moves	0.096	0.081	0.103	0.075
No changes of 1-digit industries (%)	76.27	62.49	84.98	62.49
Log weekly wages (mean)	5.90	5.94	5.73	5.83
Log weekly income (mean)	5.60	5.59	5.28	5.28
Number of observations	206,324	16,536	60,415	2,503

Note: Statistics are presented for our main sample of workers born and first observed in southern Italy. Statistics in the "North" are for those who migrated, in the "South" for those who did not. Wages and income are defined as the weekly average in a given year, they are in euros and are normalized to the 2004 sample mean.

Table 3: Baseline Specification: Model for wage

	Male			Female		
	Selection	South	North	Selection	South	North
Potential Experience (years)	-0.006	0.013	0.003	0.033	0.003	-0.004
	(0.032)	(0.000)	(0.000)	(0.000)	(0.002)	(0.798)
Potential Experience Sq.	-0.0001	0.000	0.000	-0.002	0.000	0.000
	(0.040)	(0.000)	(0.008)	(0.000)	(0.000)	(0.691)
Tenure (years)	-0.271	0.014	0.030	-0.360	0.009	0.023
	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.012)
Tenure Sq.	0.013	-0.001	-0.001	0.019	0.000	0.000
	(0.000)	(0.000)	(0.002)	(0.000)	(0.010)	(0.096)
Duration in North (years)	0.956	-0.017	0.043	1.339	0.067	-0.017
	(0.000)	(0.048)	(0.002)	(0.000)	(0.000)	(0.651)
Duration in North Sq.	-0.051	0.002	-0.002	-0.074	-0.003	0.002
	(0.000)	(0.026)	(0.004)	(0.000)	(0.034)	(0.455)
Blue collar worker	-0.016	-0.066	-0.397	0.079	-0.065	-0.326
	(0.112)	(0.000)	(0.000)	(0.008)	(0.000)	(0.000)
Manager	0.594	0.262	0.549	0.211	0.446	0.780
	(0.014)	(0.000)	(0.000)	(0.148)	(0.152)	(0.152)
Part-time worker	-0.178	0.071	-0.083	0.088	0.164	0.003
	(0.000)	(0.000)	(0.000)	(0.022)	(0.000)	(0.409)
Multi-region Firm	0.453	0.079	0.079	0.519	0.052	0.054
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.174)
Average Moves	-2.801	-	-	-5.642	-	-
	(0.000)	-	-	(0.000)	-	-
No Moves	0.071	-	-	0.144	-	-
	(0.555)	-	-	(0.168)	-	-
Inverse Mills Ratio	-	0.042	-0.023	-	-0.090	0.052
	-	(0.032)	(0.070)	-	(0.010)	(0.579)
Individual Wage Effect	-0.257	1	0.304	-0.032	1	0.300
	(0.000)	-	(0.000)	(0.002)	-	(0.000)
Constant	-1.561	5.849	5.861	-1.645	5.733	6.136
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Results are reported for the selection equation and the wage equations in the South and North of Italy. The reported coefficients are bias corrected, bootstrap p-values are reported in brackets. The variable "average moves" is the cumulative number of job changes per year in the sample, and the variable "no moves" is an indicator equal to one if the worker has never changed 1-digit industries. The "Individual Wage Effect" is the estimate of the individual fixed effect from the wage equation in Southern Italy.

Table 4: Baseline Specification: Model for Income

	Male			Female		
	Selection	South	North	Selection	South	North
Potential Experience (years)	-0.008	0.045	0.021	0.033	-0.013	0.011
	(0.936)	(0.000)	(0.000)	(0.000)	(0.002)	(0.399)
Potential Experience Sq.	0.0000	-0.001	-0.001	-0.002	0.000	0.000
	(0.046)	(0.000)	(0.000)	(0.000)	(0.000)	(0.413)
Tenure (years)	-0.231	0.087	0.104	-0.313	0.076	0.131
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tenure Sq.	0.011	-0.005	-0.006	0.017	-0.005	-0.008
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Duration in North (years)	0.938	-0.329	0.234	1.313	-0.473	0.037
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.134)
Duration in North Sq.	-0.050	0.008	-0.013	-0.072	0.005	-0.001
	(0.000)	(0.006)	(0.000)	(0.000)	(0.277)	(0.158)
Blue collar worker	-0.089	-0.098	-0.546	-0.064	-0.117	-0.484
	(0.509)	(0.000)	(0.000)	(0.776)	(0.000)	(0.000)
Manager	0.654	0.149	0.638	0.356	0.511	0.816
	(0.008)	(0.002)	(0.000)	(0.082)	(0.152)	(0.152)
Part-time worker	-0.207	-0.371	-0.713	0.026	-0.344	-0.556
	(0.000)	(0.000)	(0.000)	(0.136)	(0.000)	(0.000)
Multi-region Firm	0.493	0.082	0.208	0.587	0.020	0.164
	(0.000)	(0.000)	(0.000)	(0.000)	(0.102)	(0.126)
Average Moves	-2.496	-	-	-5.039	-	-
	(0.000)	-	-	(0.000)	-	-
No Moves	0.076	-	-	0.163	-	-
	(0.443)	-	-	(0.114)	-	-
Inverse Mills Ratio	-	0.984	-0.201	-	1.057	0.107
	-	(0.000)	(0.098)	-	(0.000)	(0.447)
Individual Income Effect	-0.362	1	0.153	-0.338	1	0.139
	(0.000)	-	(0.000)	(0.120)	-	(0.000)
Constant	-1.610	5.213	4.585	-1.753	5.430	4.953
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: Results are reported for the selection equation and the income equations in the South and North of Italy. The reported coefficients are bias corrected, bootstrap p-values are reported in brackets. The variable "average moves" is the cumulative number of job changes per year in the sample, and the variable "no moves" is an indicator equal to one if the worker has never changed 1-digit industries. The "individual income effect" is the estimate of the individual fixed effect from the income equation in Southern Italy.

Table 5: Selection of Migrants and Return Migrants

		Wage			
		Stayers in the South	Migrants	Returnees	Non-Returnees
Log-Wage	mean	5.910	5.869	5.894	5.862
	median	5.852	5.812	5.829	5.815
α	mean	0.003	-0.026	-0.015	-0.049
	median	-0.037	-0.068	-0.06	-0.08
$x'\beta$	mean	5.908	5.912	5.906	5.895
	median	5.901	5.893	5.892	5.931
		Income			
Log-Income	mean	5.615	5.682	5.544	6.835
	median	5.697	5.630	5.485	5.851
α	mean	0.007	-0.209	-0.138	-0.268
	median	0.082	-0.133	-0.085	-0.153
$x'\beta$	mean	5.607	5.892	5.672	6.104
	median	5.617	5.712	5.581	5.817

Note: We report statistics for those who never migrate ("stayers in south"), those who migrated to the North ("migrants"), those who migrated and subsequently returned to the South ("returnees"), and those who migrated but did not return ("non-returnees"). α and $x'\beta$ the value of the fixed effect and observables respectively in the outcome equation in the South.

Table 6: Monte Carlo Experiments

		Non-corrected Coefficients				Bias-corrected Coefficients			
		Mean	*Median*	*Std.Dev*	$Q_{75} - Q_{25}$	*Mean*	*Median*	*Std.Dev*	$Q_{75} - Q_{25}$
	γ	.930	.923	.174	.226	.994	.971	.234	.280
	β_S	1.9994	1.9992	.0029	.0036	1.9994	1.9992	.0030	.0035
$T = 5$	β_N	2.0093	2.0088	.0051	.008	2.0055	2.0059	.0061	.0093
	ρ	.4739	.4736	.0109	.0137	.4844	.4844	.0133	.0191
	θ	3.119	3.032	.504	.617	3.234	3.172	.677	.767
	η	5.167	5.029	.838	1.072	5.368	5.300	1.127	1.278
	γ	.935	.923	.109	.157	.980	.953	.130	.179
	β_S	1.9998	1.9999	.0020	.0022	1.9999	1.9998	.0020	.0028
$T = 10$	β_N	2.0055	2.0059	.0032	.0049	2.0028	2.0032	.0040	.0049
	ρ	.4835	.4829	.0075	.0099	.4916	.4909	.0092	.012
	θ	2.968	2.915	.299	.369	3.047	2.951	.359	.444
	η	4.932	4.844	.495	.615	5.075	4.915	.599	.737
	γ	.9588	.9557	.0889	.118	.995	.998	.106	.146
	β_S	2.0002	2.0003	.0014	.002	2.0002	2.0003	.0014	.0018
$T = 15$	β_N	2.0049	2.0048	.0024	.003	2.0027	2.0025	.003	.0036
	ρ	.4859	.4855	.0064	.0084	.4926	.4922	.0081	.0092
	θ	2.979	2.958	.237	.332	3.048	3.045	.277	.367
	η	4.955	4.913	.400	.557	5.075	5.058	.469	.654
	γ	.9628	.9636	.0778	.103	.9961	.9925	.0841	.1168
	β_S	1.9999	1.9999	.0012	.0013	1.9998	1.9999	.0012	.0013
$T = 20$	β_N	2.0046	2.0045	.0022	.0028	2.0028	2.0027	.0029	.0031
	ρ	.4872	.4875	.0051	.0062	.4926	.4928	.0072	.0104
	θ	2.973	2.961	.2216	.2958	3.037	3.008	.234	.344
	η	4.946	4.920	.365	.4073	5.060	4.987	.389	.543

Note: 100 replications of a sample of 1,000 individuals for $T = 5$, 10, 15 and 20. The values of the parameters used in the *DGP* are: $\gamma = 1$, $\beta_S = 2$, $\beta_N = 2$, $\theta = 3$, $\eta = 5$,$\rho = 0.5$. We report descriptive statistics for the sample of non-corrected coefficients in the first 4 columns and descriptive statistics for the sample of Jacknife-corrected coefficients in the last 4 columns

Table 7: Robustness: Results without Return Migrants: Model for Wage

	Male			Female		
	Selection	South	North	Selection	South	North
Potential Experience (years)	-0.003	0.013	-0.003	0.017	0.003	-0.001
	(0.000)	(0.000)	(0.000)	(0.002)	(0.008)	(0.553)
Potential Experience Sq.	0.000	0.000	0.000	-0.001	0.000	0.000
	(0.000)	(0.000)	(0.010)	(0.002)	(0.002)	(0.637)
Tenure (years)	-0.423	0.013	-0.031	-0.376	0.011	-0.024
	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)	(0.120)
Tenure Sq.	0.023	-0.001	0.003	0.021	0.000	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)	(0.046)
Blue collar worker	0.014	-0.062	-0.333	0.079	-0.065	-0.230
	(0.046)	(0.000)	(0.000)	(0.042)	(0.000)	(0.000)
Manager	0.670	0.248	0.666	-	-	-
	(0.002)	(0.000)	(0.000)	-	-	-
Part-time worker	-0.267	0.069	-0.030	0.114	0.164	0.077
	(0.000)	(0.000)	(0.012)	(0.022)	(0.000)	(0.132)
Multi-region Firm	0.823	0.074	0.181	0.874	0.049	0.157
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)
Average Moves	-1.588	-	-	-4.650	-	-
	(0.000)	-	-	(0.000)	-	-
No Moves	0.196	-	-	0.309	-	-
	(0.020)	-	-	(0.022)	-	-
Inverse Mills Ratio	-	0.042	-0.121	-	-0.013	-0.067
	-	(0.216)	(0.000)	-	(0.078)	(0.012)
Individual Wage Effect	-0.153	1	0.365	-0.069	1	0.358
	(0.433)	-	(0.000)	(0.002)	-	(0.000)
Constant	-2.125	5.844	5.846	-2.293	5.732	5.766
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note: The sample excludes return migrants. Results are reported for the selection equation and the wage equations in the south and north of Italy. The reported coefficients are bias corrected, bootstrap p-values are reported in brackets. The variable "average moves" is the cumulative number of job changes per year in the sample, and the variable "no moves" is an indicator equal to one if the worker has never changed 1-digit industries. The "Individual Wage Effect" is the estimate of the individual fixed effect from the wage equation in Southern Italy.

Table 8: Robustness: Results without Return Migrants: Model for Income

	Male			Female		
	Selection	South	North	Selection	South	North
Potential Experience (years)	-0.004	0.045	-0.018	0.016	-0.012	0.066
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.002)
Potential Experience Sq.	0.000	-0.001	0.000	-0.001	0.000	-0.002
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.004)
Tenure (years)	-0.375	0.078	-0.304	-0.331	0.086	-0.358
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tenure Sq.	0.020	-0.005	0.018	0.019	-0.005	0.023
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Blue collar worker	-0.067	-0.102	-0.596	-0.040	-0.116	-0.368
	(0.645)	(0.000)	(0.000)	(0.723)	(0.000)	(0.036)
Manager	0.740	0.171	1.223	-	-	-
	(0.000)	(0.000)	(0.000)	-	-	-
Part-time worker	-0.302	-0.375	-0.917	0.061	-0.355	-0.536
	(0.000)	(0.000)	(0.000)	(0.092)	(0.000)	(0.072)
Multi-region Firm	0.869	0.064	1.211	0.908	-0.097	1.345
	(0.000)	(0.000)	(0.000)	(0.000)	(0.154)	(0.000)
Average Moves	-1.715	-	-	-4.345	-	-
	(0.000)	-	-	(0.000)	-	-
No Moves	0.205	-	-	0.312	-	-
	(0.010)	-	-	(0.018)	-	-
Inverse Mills Ratio	-	0.942	-1.203	-	2.469	-1.355
	-	(0.058)	(0.000)	-	(0.000)	(0.000)
Individual Income Effect	-0.334	1	0.081	-0.285	1	-0.114
	(0.000)	-	(0.030)	(0.469)	-	(0.010)
Constant	-2.099	5.240	2.901	-2.308	5.425	1.907
	(0.000)	(0.000)	(0.998)	(0.000)	(0.000)	(0.968)

Note: Sample excludes return migrants. Results are reported for the selection equation and the income equations in the south and north of Italy. The reported coefficients are bias corrected, bootstrap p-values are reported in brackets. The variable "average moves" is the cumulative number of job changes per year in the sample, and the variable "no moves" is an indicator equal to one if the worker has never changed 1-digit industries. The "individual income effect" is the estimate of the individual fixed effect from the income equation in Southern Italy.

Table 9: Robustness: Results without Central Region: Model for Wage

a. Selection Equation (Independent variable = 1 if worker observed in North)

	Male, All		Male, No Returnees		Female, All		Female, No Returnees	
Duration in North (years)	0.955	(0.000)	-	-	1.341	(0.000)	-	-
Duration in North Sq.	-0.052	(0.000)	-	-	-0.073	(0.000)	-	-
Potential Experience (years)	-0.010	(0.822)	-0.002	(0.000)	0.039	(0.000)	0.023	(0.006)
Potential Experience Sq.	0.000	(0.110)	0.000	(0.000)	-0.002	(0.000)	-0.001	(0.000)
Blue collar worker	0.022	(0.006)	0.061	(0.004)	0.146	(0.006)	0.162	(0.020)
Average Moves	-2.829	(0.000)	-1.613	(0.000)	-5.947	(0.000)	-5.149	(0.000)
No Moves	0.056	(0.743)	0.180	(0.178)	0.174	(0.144)	0.384	(0.012)
Individual Wage Effect	-0.287	(0.000)	-0.128	(0.299)	-0.063	(0.036)	-0.050	(0.006)

b. Wage Equation in South

	Male, All		Male, No Returnees		Female, All		Female, No Returnees	
Potential Experience (years)	0.013	(0.000)	0.000	(0.000)	0.003	(0.002)	0.003	(0.012)
Potential Experience Sq.	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.002)
Duration in North (years)	-0.014	(0.108)	-	-	0.058	(0.002)	-	-
Duration in North Sq.	0.002	(0.030)	-	-	-0.002	(0.052)	-	-
Blue collar worker	-0.066	(0.000)	-0.062	(0.000)	-0.065	(0.000)	-0.065	(0.000)
Inverse Mills Ratio	0.038	(0.136)	0.065	(0.176)	-0.086	(0.022)	-0.023	(0.056)
Individual Wage Effect	1	-	1	-	1	-	1	-

c. Wage Equation in North

	Male, All		Male, No Returnees		Female, All		Female, No Returnees	
Potential Experience (years)	0.004	(0.002)	-0.006	(0.004)	-0.011	(0.617)	0.001	(0.281)
Potential Experience Sq.	0.000	(0.006)	0.000	(0.016)	0.000	(0.535)	0.000	(0.164)
Duration in North (years)	0.041	(0.014)	-	-	-0.053	(0.481)	-	-
Duration in North Sq.	-0.002	(0.052)	-	-	0.004	(0.295)	-	-
Blue collar worker	-0.403	(0.000)	-0.330	(0.000)	-0.322	(0.000)	-0.180	(0.000)
Inverse Mills Ratio	-0.016	(0.255)	-0.057	(0.000)	0.101	(0.337)	0.021	(0.026)
Individual Wage Effect	0.281	(0.000)	0.339	(0.000)	0.301	(0.000)	0.347	(0.002)

Note: Sample excludes the central Italian provinces of Lazio, Marche and Umbria. Results are reported for the selection equation and the wage equations in the south and north of Italy. The reported coefficients are bias corrected, bootstrap p-values are reported in brackets. The variable "average moves" is the cumulative number of job changes per year in the sample, and the variable "no moves" is an indicator equal to one if the worker has never changed 1-digit industries. The "Individual Wage Effect" is the estimate of the individual fixed effect from the wage equation in Southern Italy.

Table 10: Robustness: Results without Central Region: Model for Income

a. Selection Equation (Independent variable = 1 if worker observed in North)

	Male, All		Male, No Returnees		Female, All		Female, No Returnees	
Duration in North (years)	0.936	(0.000)	-	-	1.311	(0.000)	-	-
Duration in North Sq.	-0.052	(0.000)	-	-	-0.071	(0.000)	-	-
Potential Experience (years)	-0.012	(0.764)	-0.003	(0.000)	0.037	(0.000)	0.022	(0.004)
Potential Experience Sq.	0.000	(0.144)	0.000	(0.000)	-0.002	(0.000)	-0.001	(0.000)
Blue collar worker	-0.057	(0.094)	-0.028	(0.044)	-0.006	(0.150)	0.042	(0.146)
Average Moves	-2.451	(0.000)	-1.719	(0.000)	-5.271	(0.000)	-4.775	(0.000)
No Moves	0.064	(0.639)	0.190	(0.128)	0.194	(0.092)	0.389	(0.010)
Individual Income Effect	-0.402	(0.000)	-0.345	(0.000)	-0.359	(0.058)	-0.276	(0.305)

b. Income Equation in South

	Male, All		Male, No Returnees		Female, All		Female, No Returnees	
Potential Experience (years)	0.045	(0.000)	0.045	(0.000)	-0.010	(0.016)	-0.011	(0.002)
Potential Experience Sq.	-0.001	(0.000)	-0.001	(0.000)	0.000	(0.000)	0.000	(0.000)
Duration in North (years)	-0.278	(0.000)	-	-	-0.303	(0.030)	-	-
Duration in North Sq.	0.008	(0.022)	-	-	-0.001	(0.204)	-	-
Blue collar worker	-0.102	(0.000)	-0.105	(0.000)	-0.121	(0.000)	-0.124	(0.000)
Inverse Mills Ratio	0.920	(0.000)	0.887	(0.653)	0.830	(0.000)	2.279	(0.000)
Individual Income Effect	1	-	1	-	1	-	1	-

c. Income Equation in North

	Male, All		Male, No Returnees		Female, All		Female, No Returnees	
Potential Experience (years)	0.017	(0.000)	-0.020	(0.000)	0.017	(0.228)	0.080	(0.016)
Potential Experience Sq.	-0.001	(0.000)	0.000	(0.000)	0.000	(0.293)	-0.003	(0.018)
Duration in North (years)	0.254	(0.000)	-	-	0.027	(0.054)	-	-
Duration in North Sq.	-0.014	(0.000)	-	-	0.000	(0.096)	-	-
Blue collar worker	-0.569	(0.000)	-0.573	(0.014)	-0.490	(0.000)	-0.279	(0.146)
Inverse Mills Ratio	-0.214	(0.042)	-1.399	(0.000)	0.110	(0.687)	-1.037	(0.000)
Individual Income Effect	0.128	(0.000)	-0.030	(0.351)	0.096	(0.000)	-0.074	(0.070)

Note: Sample excludes the central Italian provinces of Lazio, Marche and Umbria. Results are reported for the selection equation and the income equations in the south and north of Italy. The reported coefficients are bias corrected, bootstrap p-values are reported in brackets. The variable "average moves" is the cumulative number of job changes per year in the sample, and the variable "no moves" is an indicator equal to one if the worker has never changed 1-digit industries. The "individual income effect" is the estimate of the individual fixed effect from the income equation in Southern Italy.

Table 11: Regression of Residuals from Mincerian Wage Regressions on Education Variables

VARIABLES	Residual from	Residual from 2004-2007 Panel
	2004 OLS Regression	Fixed Effect Regression
Upper secondary and	0.058***	0.026
post secondary education	-0.012	-0.018
Tertiary education	0.18***	0.063*
	-0.02	-0.036
Constant	-0.054***	-0.024*
	-0.0091	-0.013
Observations	5,957	13,713
R-squared	0.025	0.000

Note: Estimated from IT-SILC data. Reference group are workers with lower secondary education or less. Robust standard errors in parentheses. *** p¡0.01, ** p¡0.05, * p¡0.1. First stage regressions include gender, occupation, industry location and firm size dummies, age and experience and those squared.