



CENTRE D'ETUDES
ET DE RECHERCHES
SUR LE DEVELOPPEMENT
INTERNATIONAL

SERIE ETUDES ET DOCUMENTS DU CERDI

Can selective immigration policies reduce migrants' quality?

Simone BERTOLI

Vianney DEQUIEDT

Yves ZENOU

Etudes et Documents n° 19

September 2014

CERDI
65 BD. F. MITTERRAND
63000 CLERMONT FERRAND - FRANCE
TEL. 04 73 17 74 00
FAX 04 73 17 74 28
www.cerdi.org

The authors

Simone BERTOLI, Clermont Université, Université d'Auvergne, CNRS, UMR 6587, CERDI, F-63009 Clermont Fd
Email: simone.bertoli@udamail.fr

Vianney DEQUIEDT, Clermont Université, Université d'Auvergne, CNRS, UMR 6587, CERDI, F-63009 Clermont Fd
Email: vianney.dequiedt@udamail.fr

Yves ZENOU, Stockholm University, Department of Economics, 106 91 Stockholm, Sweden
Email: yves.zenou@ne.su.se

Corresponding author: Simone Bertoli



This work was supported by the LABEX IDGM+ (ANR-10-LABX-14-01) within the program “Investissements d’Avenir” operated by the French National Research Agency (ANR)

La série des *Etudes et Documents* du CERDI est consultable sur le site :
<http://www.cerdi.org/ed>

Directeur de la publication : Vianney Dequiedt
Directeur de la rédaction : Catherine Araujo Bonjean
Responsable d'édition : Annie Cohade
ISSN : 2114 - 7957

Avertissement :

Les commentaires et analyses développés n'engagent que leurs auteurs qui restent seuls responsables des erreurs et insuffisances.

Abstract

Destination countries have been resorting to selective immigration policies to improve migrants' quality. We propose a model that analyzes the effects of selective immigration policies on migrants' quality, measured by their wages at destination. Screening potential migrants on the basis of observable characteristics also influences their self-selection on unobservables that influences their wages. We show that the prevailing pattern of selection on unobservables influences the effect of an increase in selectivity, which can reduce migrants' quality when migrants are positively self-selected.

Mots clés / Key Words: selective policies; self-selection; migrants' quality.

Codes JEL / JEL classification: F22; K37; J61.

Acknowledgements / Remerciements

The authors are grateful to Jesús Fernández-Huertas Moraga and to the other participants of the Seventh Migration and Development Conference for their comments and suggestions. Simone Bertoli and Vianney Dequiedt thank the FERDI and the *Agence Nationale de la Recherche* of the French government through the program "Investissements d'avenir" (ANR-10-LABX-14-01) for their support. The usual disclaimers apply.

Can selective immigration policies reduce migrants' quality?*

Simone Bertoli^a, Vianney Dequiedt^b, and Yves Zenou^c

^a*University of Auvergne[†] and CNRS*

^b*University of Auvergne[‡]*

^c*Stockholm University and IFN[§]*

September 15, 2014

Abstract

Destination countries have been resorting to selective immigration policies to improve migrants' quality. We propose a model that analyzes the effects of selective immigration policies on migrants' quality, measured by their wages at destination. Screening potential migrants on the basis of observable characteristics also influences their self-selection on unobservables that influences their wages. We show that the prevailing pattern of selection on unobservables influences the effect of an increase in selectivity, which can reduce migrants' quality when migrants are positively self-selected.

Keywords: selective policies; self-selection; migrants' quality.

JEL codes: F22; K37; J61.

*The authors are grateful to Jesús Fernández-Huertas Moraga and to the other participants at the Seventh Migration and Development Conference for their comments and suggestions. Simone Bertoli and Vianney Dequiedt thank the FERDI and the *Agence Nationale de la Recherche* of the French government through the program "*Investissements d'avenir*" (ANR-10-LABX-14-01) for their support. The usual disclaimers apply.

[†]CERDI, Bd. F. Mitterrand, 65, F-63000, Clermont-Ferrand. Email: simone.bertoli@udamail.fr .

[‡]CERDI, Bd. F. Mitterrand, 65, F-63000, Clermont-Ferrand. Email: vianney.dequiedt@udamail.fr.

[§]Department of Economics, 106 91 Stockholm, Sweden. Email: yves.zenou@ne.su.se.

“Remarkably little is known about [...] whether the chosen policy, in fact, has the desired outcomes in terms of the size and composition of the immigrant flow.”

George J. Borjas (2014), *Immigration Economics*

1 Introduction

Destination countries are deeply concerned about the composition and scale of incoming migration flows as they contribute to shape both the overall economic impact of immigration and its distributional effects. The economic literature has traditionally relied on market prices to measure immigrants’ quality through their earnings upon arrival at destination, and evidence of a fall in migrants’ initial earnings in recent decades¹ has prompted debates around the need to reform immigration policies in order to reverse this declining trend.² Specifically, a growing number of countries are moving towards immigration policies that screen potential immigrants on the basis of their observable characteristics, such as education and language proficiency (Bertoli *et al.*, 2012), granting better chances of admission at destination to applicants endowed with more desirable individual characteristics.

While the (narrow sets of) characteristics upon which potential migrants are selected are related to their earnings at destination, it is important to acknowledge that some other relevant determinants of migrants’ quality, such as ability and motivation, remain *unobservable* for the immigration officers. These unobservable characteristics can enter into the decision to self-select into migration (Roy, 1951; Borjas, 1987), so that the effectiveness of selective immigration policies in raising migrants’ quality also depends on how they influence the pattern of self-selection on unobservables. The possible impact of the out-selection mechanisms adopted by the countries of destination on the prevailing pattern of selection on unobservables contributes to shape the ultimate effect of the immigration policy, as “education accounts for only a small portion of the variance in earnings across workers, suggesting that the nature of selection in educational attainment may not necessarily “transfer over” to a more comprehensive measure of a worker’s human capital” (Borjas, 2014, pp. 29-30).³

¹See, for instance, Borjas (1985) and Borjas and Friedberg (2009) for the United States, and Aydemir and Skuterud (2005) for Canada.

²“Most discussions of immigration policy “run” with one of the facts about the economic impact of immigration—that immigrants reduce the wage of native workers, or that more recent immigrants tend to be relatively less skilled—to propose some type of reform in immigration policy.” (Borjas, 1999a, p. 182).

³Along the same lines, Kaestner and Malamud (2014) caution about the limits of “using individual

For instance, the analysis by Aydemir (2011) reveals that, as expected, the Canadian points system effectively increases the average level of migrants' education but that "immigrants admitted for their skills do *not* necessarily perform better in the labor market." (Aydemir, 2011, p. 451, emphasis added).^{4,5} This, in turn, suggests that a focus on observable skills can produce only a partial, and possibly misleading, account of the effects of selective immigration policies on migrants' quality.

This paper proposes an extension of the seminal paper by Borjas (1987) to analyze how selective immigration policies influence migrants' quality when migrants are self-selected on unobservables related to the earnings at destination. Specifically, we consider a two-country model where potential migrants are heterogeneous with respect to both education and ability and where the destination country imposes higher policy-induced migration costs on uneducated potential migrants. We analyze the effect on migrants' quality of a scale-preserving increase in selectivity,⁶ which is defined as a reduction of migration costs for educated applicants, matched by a simultaneous increase in moving costs for uneducated ones that leaves the total scale of incoming migration flows unchanged.

The analysis reveals that the response of migrants' quality to a scale-preserving increase in selectivity hinges on the prevailing pattern of selection on ability. When immigrants are positively selected on ability, so that migrants' average (log) wage at destination exceeds the corresponding (hypothetical) average wage of the non-migrants with identical observable characteristics, then a scale-preserving increase in selectivity can reduce migrants' quality when selectivity is pushed too far. This occurs because the direct beneficial effect of the policy change is thwarted by an opposite negative effect, due to the induced reduction in the average wage of the educated migrants. We demonstrate that there is an optimal degree of

components of skill such as education to assess migrant selection with respect to skill" (p. 89).

⁴Antecol *et al.* (2003) question the ability of the Canadian immigration policy to improve migrants' observable characteristics, as compared to the United States, using data from the 1991 Canadian population census.

⁵Ambrosini and Peri (2012) find that the lower earnings of Mexican migrants to the United States with respect to stayers are "mostly due to [selection] on unobserved wage-earning characteristics and not on observed ones" (p. 147), while Fernández-Huertas Moraga (2011) and Kaestner and Malamud (2014) find that a larger role is played by observables, with this latter paper also including measures of cognitive ability among the observable characteristics.

⁶This is similar in spirit to Biavaschi and Elsner (2013) who analyze the welfare implications for the sending and the receiving countries of a change in the pattern of migrants' selection for a constant scale of migration flows.

selectivity in immigration policies when migrants are positively selected on unobservables, and that further increases in selectivity are detrimental to migrants' quality. No such a perverse effect arises when the opposite pattern of selection on unobservables prevails, as the direct positive effect of the policy change is reinforced by the ensuing increase in the average wage of educated migrants.

This theoretical result is robust with respect to several extensions of the basic version of the model. Specifically, we analyze the implications of (i) introducing general equilibrium effects induced by migration, (ii) introducing unobserved heterogeneity in the preferences for migration, (iii) considering that wages are only *locally* observable, and (iv) allowing for a change in the informational structure of educated agents.

The forces at play in our theoretical model are related to the ones analyzed by Bertoli and Rapoport (2015). In that paper, the effect of an expansion of the size of migration networks on migrants' selection on education depends on the endogeneity of the distribution of education at origin with respect to variations in the prospect to migrate. The emphasis put on the potentially perverse effect of selectivity on observables is reminiscent of results in the moral-hazard multitasking literature (Holmstrom and Milgrom, 1991). There, it is a well-known result that designing high-powered incentive schemes on easily observable tasks may lead the agent to divert effort from tasks that are more difficult to monitor and may *in fine* hurt the principal. The same logic applies here to the different dimensions of migrants' quality.

This paper is mainly related to two strands of literature. First, it is related to the literature on migrants' selection (Borjas, 1987; Antecol *et al.*, 2003; Chiquiar and Hanson, 2005; Jasso and Rosenzweig, 2009; Fernández-Huertas Moraga, 2011, 2013; Ambrosini and Peri, 2012; Dequiedt and Zenou, 2013; Kaestner and Malamud, 2014), including the papers that analyze the determinants of selection on education (McKenzie and Rapoport, 2010; Bertoli, 2010a; Beine *et al.*, 2011). Second, it is also related to the papers that analyze the influence of immigration policies on migrants' selection on education, both from a theoretical (Bellettini and Berti Ceroni, 2007; Docquier *et al.*, 2008; Bertoli and Brücker, 2011; Bianchi, 2013; Bertoli and Rapoport, 2015) and an empirical perspective (Antecol *et al.*, 2003; Jasso and Rosenzweig, 2009; Aydemir, 2011; Belot and Hatton, 2012).

The rest of the paper is structured as follows: Section 2 introduces our model. Section 3 analyzes the effects of selective immigration policies on migrants' quality in a basic version

of our theoretical model, and Section 4 discusses the implications of a number of extensions. Finally, Section 5 concludes.

2 The model

We develop a random utility maximization model describing the location-decision problem that potential migrants face. That model largely follows and extends Borjas (1987), which is derived from the seminal contribution by Heckman (1979). Consider an origin country, which is denoted with the subscript 0, with a population of mass one of agents, which are indexed by i . We assume that the origin country's population can be either educated (e) or uneducated (u), with $\alpha \in (0, 1)$ denoting the exogenous share of educated agents. Agents can choose between a domestic job in country 0 and a foreign job in country 1. Education is an observable characteristic in both countries and influences agents' wage. The expected values of the domestic and foreign log wages verify $\mu_j^e - \mu_j^u > 0$ for $j = 0, 1$.⁷

Other individual characteristics remain unobservable but also influence agents' productivity in both countries. Specifically, we assume that wages in both countries and for both education levels follow a log-normal distribution. More precisely, for agent i , with education level l , in country j , we have:

$$\ln(w_{ij}^l) = \mu_j^l + \epsilon_{ij},$$

with $j = 0, 1$ and $l \in \{e, u\}$, and⁸

⁷Notice that $\mu_0^e - \mu_0^u$ and $\mu_1^e - \mu_1^u$ should not be interpreted as representing the causal impact of education on the (log) wages in the two countries, as educated agents might be non-randomly self-selected on unobservables; the two terms simply represent the differences in the average log wage for educated and uneducated agents when all of them are employed in either of the two countries, so that these differences can also pick up differences in unobservables between the two group of agents. Endogenizing the choice of education goes beyond the scope of the current paper.

⁸Borjas (1987) proposes a static model, which portrays migration as a one-time permanent decision. Borjas (1999b) observes that (1) could be interpreted "as giving the distributions of the present value of the earnings stream in each country", as in Sjaastad (1962). The model can also be interpreted as describing a sequence of repeated location choices, with time-invariant distributions for wages and no serial correlation in the realizations of the stochastic components (Fernández-Huertas Moraga, 2013).

$$\begin{pmatrix} \ln(w_{i0}^l) \\ \ln(w_{i1}^l) \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\mu}^l, \Sigma), \quad (1)$$

where

$$\boldsymbol{\mu}^l = (\mu_0^l, \mu_1^l)' \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}.$$

We assume for simplicity that the covariance matrix of the bivariate normal distribution of log wages is the same for educated and uneducated agents. The distribution of wages in the two countries is assumed to be exogenous, so that migration does not produce general equilibrium effects.

For agent i , opting for a foreign job requires paying a migration cost whose monetary equivalent stands at C_i and which may include both pecuniary and non-pecuniary costs, such as the psychological costs of being away from home. As in Borjas (1987), we assume that the time-equivalent migration costs, defined as the ratio between C_i and the individual-specific wage at origin w_{i0}^l , do *not* vary across individuals with the same level of education.⁹ This implies that self-selection into migration is driven exclusively by observable and unobservable factors that influence the wages in the two countries, while agents are not heterogeneous in their preferences for migration due to non-wage factors.¹⁰

Wages are assumed to be *remotely* observable, so that agents decide whether to migrate or not after observing the realizations of the stochastic component of domestic and foreign wages.¹¹ For an individual with level of education $l \in \{e, u\}$, migration represents a utility-maximizing decision if and only if:¹²

$$\ln(w_{i0}^l) + \pi^l < \ln(w_{i1}^l),$$

⁹This implies that agents are heterogeneous with respect to the monetary equivalent of migration costs, with C_i following a log-normal distribution that is perfectly correlated with w_{i0}^l .

¹⁰The assumption of the invariance of time-equivalent migration costs is relaxed in Borjas (1999b); see also Section 4.2 below.

¹¹Bertoli (2010b) considers an alternative informational structure, with wages that are *locally* observable, so that the decision to migrate is taken before observing the realization of $\ln(w_{i1})$. Implications of such an alternative information structure are analyzed in Section 4.3 below.

¹²Borjas (1987) relies on the approximation, $\ln(1 + C_i/w_{i0}) \approx C_i/w_{i0}$, which is accurate only when C_i is sufficiently close to zero, but the analysis of the whole model does not hinge on this approximation that we do not retain here.

where $\pi^l = \ln(1 + C_i/w_{i0}^l)$.¹³ Educated and uneducated agents face different time-equivalent migration costs. Specifically, we follow the literature (Schultz, 1975; Chiquiar and Hanson, 2005; McKenzie and Rapoport, 2010; Beine *et al.*, 2011) by assuming that $\pi^e < \pi^u$. The inequality above, which describes self-selection into migration, can be rewritten as:

$$\epsilon_{i2} \equiv \epsilon_{i1} - \epsilon_{i0} > \mu_0^l + \pi^l - \mu_1^l. \quad (2)$$

The probability that (2) is satisfied is given by:

$$\Pr(\epsilon_{i2} > \mu_0^l + \pi^l - \mu_1^l) = 1 - \Phi(z^l), \quad (3)$$

where $\Phi(\cdot)$ represents the cumulative distribution of a standard normal and where

$$z^l = \frac{\mu_0^l + \pi^l - \mu_1^l}{\sigma_2},$$

with σ_2 being the standard deviation of the differential ϵ_2 between the two stochastic components of the (log) wages. A (nearly) universal empirical regularity is that the propensity to migrate is higher among individuals with tertiary education than among less-educated individuals (see, for instance, Docquier *et al.*, 2009). In terms of our model, this requires that $z^e < z^u$. This condition does *not* require that the return to education is larger at destination than at origin, i.e., $\mu_1^e - \mu_1^u > \mu_0^e - \mu_0^u$, as a greater propensity to migrate among educated agents could be induced by lower time-equivalent migration costs that they face.

Migrants represent a self-selected portion of the population at origin, so that the conditional expectation of $\ln(w_{i1}^l)$ among the migrants in general differs from the unconditional expected value μ_1^l . The assumption of bivariate normality implies that (Heckman, 1979; Borjas, 1987):

$$E[\ln(w_{i1}^l) | \epsilon_{i2} > z^l] = \mu_1^l + \gamma\lambda(z^l), \quad (4)$$

where γ is given by the covariance between the conditioning variable ϵ_2 , which drives the decision to self-select into migration, and the stochastic component ϵ_1 of $\ln(w_{i1}^l)$, scaled by the standard deviation of the conditioning variable:

¹³With a minor abuse of terminology, we will be referring to π as time-equivalent migration costs.

$$\gamma = \frac{\sigma_{12}}{\sigma_2} = \frac{\sigma_1^2 - \sigma_{01}}{(\sigma_1^2 + \sigma_0^2 - 2\sigma_{01})^{1/2}}$$

and where $\lambda(\cdot)$ represents the Inverse Mills ratio:

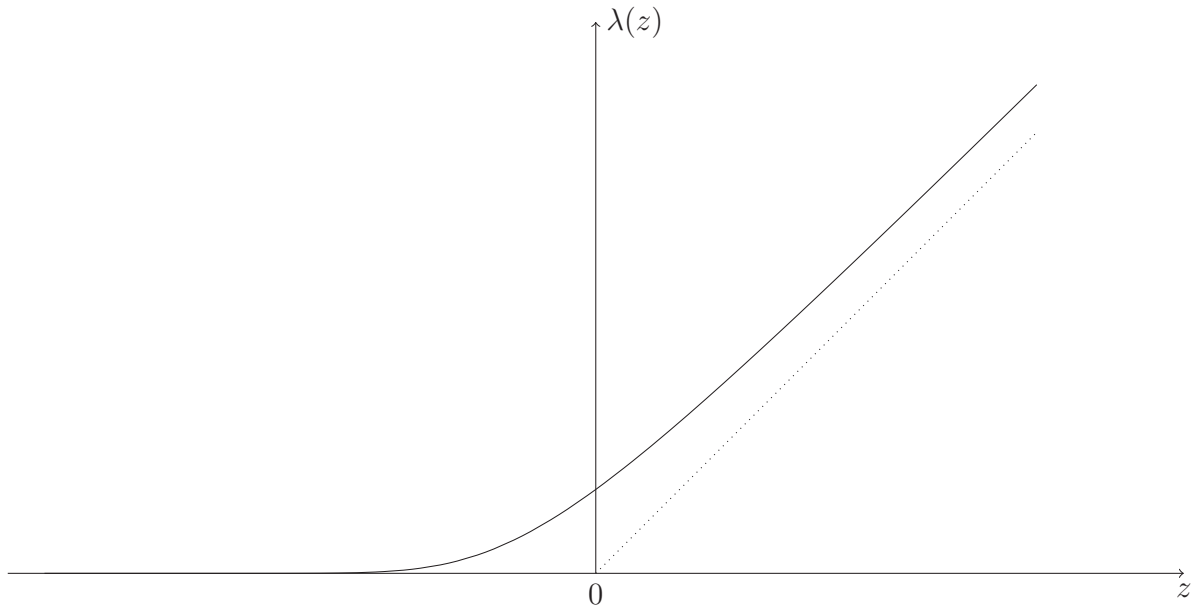
$$\lambda(z) = \frac{\phi(z)}{1 - \Phi(z)},$$

with $\phi(z)$ and $\Phi(z)$ being the density function and the CDF of the standard normal distribution. The Inverse Mills ratio $\lambda(z)$ gives the expected value of the upper tail of a standard normal distribution truncated at z , and it is thus a positive and increasing function of z , as shown in Figure 1. Furthermore, it can be shown that (Heckman, 1979, p. 157):¹⁴

$$\frac{\partial \lambda(z)}{\partial z} = \lambda(z)[\lambda(z) - z] < 1, \tag{5}$$

so that $\lambda(\cdot)$ is a contraction mapping, a property that will be used in our analysis below.

Figure 1: The Inverse Mills ratio



¹⁴The proof of the inequality in (5) indirectly follows from the fact that the variance of the upper tail of the truncated standard normal distribution is equal to $1 - \partial \lambda(z) / \partial z$.

Borjas (1987) defines Q_1 as the difference between the expected value of the distribution of $\ln(w_1)$ over the sample of migrants and the expected value of the entire distribution. Adapted to our context with two levels of education, we have for $l \in \{e, u\}$:

$$Q_1^l \equiv E [\ln(w_{i1}^l) | \epsilon_{i2} > z^l] - E [\ln(w_{i1}^l)] = \gamma \lambda(z^l). \quad (6)$$

We say that migrants with level of education l are positively (negatively) selected on unobservables if $Q_1^l > 0$ ($Q_1^l < 0$). Remarkably, the sign of (6), and hence the pattern of selection on unobservables, depends exclusively on γ . The assumption that the covariance matrix Σ does not vary with l implies that the same pattern of selection on unobservables characterizes both educated and uneducated migrants. However, the deterministic components of the log wages μ_0^l and μ_1^l and time-equivalent migration costs π^l determine the extent to which the expected value of the foreign wages for the migrants differ from the *unconditional* expected value; therefore this extent, related to the intensity of selection (Borjas, 2014), may differ for the two groups. Letting the coefficient of correlation between ϵ_0 and ϵ_1 be denoted by ρ_{01} , we can rewrite γ as follows:

$$\gamma = \frac{\sigma_1^2 - \sigma_{01}}{\sigma_2} = \frac{\sigma_0 \sigma_1}{\sigma_2} \left(\frac{\sigma_1}{\sigma_0} - \rho_{01} \right).$$

We immediately see that $\sigma_1 > \sigma_0$ represents a sufficient condition to have a pattern of positive selection on unobservables, as $|\rho_{01}| \leq 1$.

Furthermore, (6) and the fact that $\lambda(z^l)$ is a monotonically increasing function of z^l imply that a marginal variation in z^l increases (reduces) Q_1^l when migrants are positively (negatively) selected on unobservables. As z^l is positively related to time-equivalent migration costs π^l , this implies that a *reduction* in migration costs always dilutes the extent of migrants' non-random (either positive or negative) selection on unobservables. Formally, we have that:

$$\frac{\partial |Q_1^l|}{\partial \pi} = |\gamma| \frac{[\lambda(z^l) - z^l] \lambda(z^l)}{\sigma_2} > 0. \quad (7)$$

This occurs because a reduction in π^l increases the probability of migrating $1 - \Phi(z^l)$, thus reducing the distance between the unconditional and the conditional expected value of the distribution of $\ln(w_1^l)$, and hence the intensity of selection for group l .

3 Selective immigration policies and migrants' quality

As discussed in Section 2, migration costs are, at least partly, policy-induced by the recipient country through the legal framework that regulates immigrants' admission at destination. A number of papers have modeled the influence of immigration policies on migration decisions in terms of the monetary costs that they, implicitly or explicitly, impose. See, for example, Giordani and Ruta (2013), Bianchi (2013) and Docquier *et al.* (2015).^{15,16}

The random allocation of a fixed number of immigration visas through a lottery among the applicants is an alternative way of modeling immigration policies. This allows representing selectivity through a variation in the probabilities of success in the lottery for different groups of applicants (see Mountford, 1997, Beine *et al.*, 2001, Bertoli and Brücker, 2011 or Bertoli and Rapoport, 2015). This type of selective immigration policy will not alter the predictions of our model as long as there is a cost in participating in the migration lottery.¹⁷

Destination countries can impose different migration costs on potential migrants with different observable characteristics, such as education. We assume that destination countries have (possibly different) preferences both over the scale of immigration and over migrants' quality. As in Bertoli and Brücker (2011), we do not impose any structure on the trade-off between migrants' quantity and quality and we just assume that the objective function of the destination country is increasing with migrants' quality for any given scale of migration. Migrants' quality depends both on their observable and *unobservable* characteristics. Market prices provide a natural way of combining the effects of both types of characteristics and we follow the literature by defining quality as the average log wage that migrants earn at destination.¹⁸ Thus, letting n and y denote the scale and quality of migration, we just assume that $(n, y_2) \succ (n, y_1)$ if $y_2 > y_1$.¹⁹

¹⁵Grogger and Hanson (2011) and Bertoli *et al.* (2013) recover the implicit migration costs that reconcile observed migration flows with utility-maximizing destination choices.

¹⁶Sending countries can also impose policy-induced migration costs on their citizens, with the so-called Mariel boatlift representing a famous instance of the effects of relaxing the barriers to emigration (Card, 1990). See also McKenzie (2007) for empirical evidence on the relevance of the costs of obtaining a passport.

¹⁷If participation in the lottery is costless, then out-selection mechanisms have no influence on self-selection decisions.

¹⁸See, for instance, Borjas (1985) or Aydemir (2011).

¹⁹“[T]he relative skills of immigrants determine the economic benefits from immigration. The United States benefits from international trade because it can import goods that are not available or are too expensive to produce in the domestic market. Similarly, the country benefits from immigration because it

From (3), the scale of migration is given by:

$$n(\pi^e, \pi^u) = \alpha[1 - \Phi(z^e)] + (1 - \alpha)[1 - \Phi(z^u)], \quad (8)$$

where the notation is meant to emphasize the dependence of n on time-equivalent migration costs. A change in time-equivalent migration costs π^e and π^u is scale-invariant if it does not change the value of $n(\pi^e, \pi^u)$ in (8). By totally differentiating (8), we see that a scale-invariant change in migration costs satisfies:

$$\frac{\partial \pi^e}{\partial \pi^u} = \frac{\partial z^e}{\partial z^u} = -\frac{(1 - \alpha)}{\alpha} \frac{\phi(z^u)}{\phi(z^e)} < 0. \quad (9)$$

Let $f_k(\pi^u)$ represents the function, which is implicitly defined by (9), that provides us with (if it exists) the unique value of π^e , which gives rise to a scale of migration $n(\pi^e, \pi^u) = k$ when time-equivalent migration costs for uneducated individuals stand at π^u . Formally, we define a scale-preserving increase in selectivity as an increase in π^u and a decrease in π^e satisfying $\pi^e = f_k(\pi^u)$.

According to our definition, migrants' quality is simply a weighted average of the conditional log wages for the two types of migrants:

$$y(\pi^e, \pi^u) = s^e [\mu_1^e + \gamma\lambda(z^e)] + (1 - s^e) [\mu_1^u + \gamma\lambda(z^u)], \quad (10)$$

where the share of educated migrants is given by $s^e = \alpha[1 - \Phi(z^e)]/n(\pi^e, \pi^u)$.²⁰

A scale-preserving increase in selectivity influences migrants' quality $y(\pi^e, \pi^u)$ through two distinct channels: (i) it increases the share s^e of educated agents among the migrants, whose log wages are drawn from a distribution with a higher expected value $\mu_1^e > \mu_1^u$, and (ii) it modifies the intensity of selection for both educated and uneducated migrants. While (i) directly improves migrants' quality,²¹ the influence of (ii) on quality can go in either direction, and we demonstrate that a scale-preserving increase in selectivity can determine

can import workers with scarce qualifications and abilities." (Borjas and Friedberg, 2009, p. 3).

²⁰Notice that a scale-preserving increase in selectivity could be equivalently defined as an increase in s^e , for a constant $n(\pi^e, \pi^u)$.

²¹We are implicitly ruling out here the case, which is theoretically possible but empirically uninteresting, where educated migrants earn less than uneducated migrants.

an adverse effect on educated migrants' selection on unobservables that offsets the positive direct effect of the policy change, thus reducing migrants' quality.

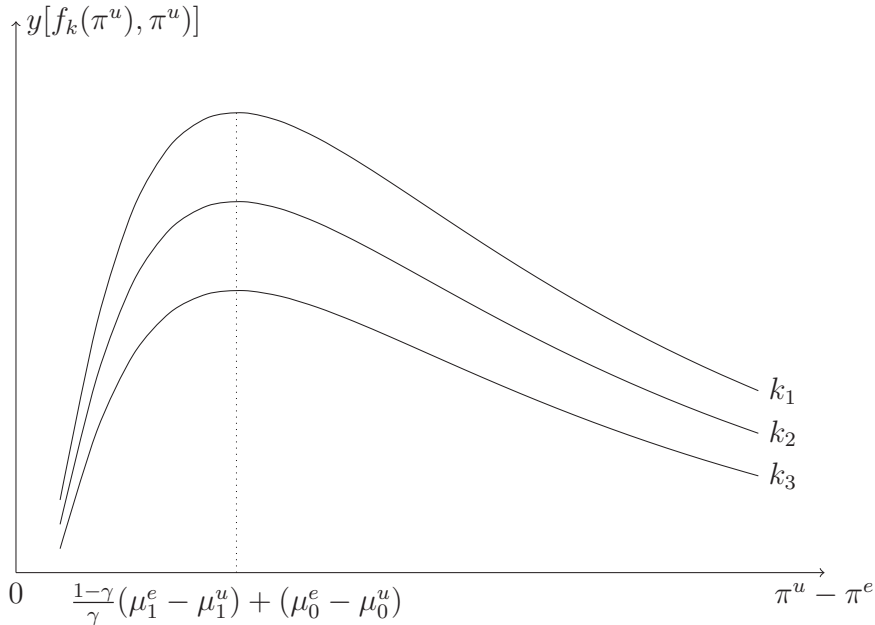
The following proposition contains the central theoretical prediction of our model:

Proposition 1 *If migrants are positively selected on unobservables, then a scale-preserving increase in selectivity reduces migrants' quality if and only if $\pi^u - \pi^e > \frac{1-\gamma}{\gamma}(\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$.*

Proof. See Appendix A.1.

Proposition 1 identifies the differential between π^u and π^e that maximizes migrants' quality for a given scale of migration. When migrants are positively selected on unobservables, then a scale-invariant increase in selectivity that pushes $\pi^u - \pi^e$ above the quality-maximizing value of $\frac{1-\gamma}{\gamma}(\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$ reduces migrants' quality, as shown in Figure 2.

Figure 2: Migrants' quality $y[f_k(\pi^u), \pi^u]$ and $\pi^u - \pi^e$



Note: the figure represents the evolution of migrants' quality along three iso-migration curves with $k_1 < k_2 < k_3$ when $\gamma > 0$.

While the patterns of selection on education and on unobservables can differ, Borjas (2014) observes that it is unclear “why the relative rates of return to skills between any

two countries (which presumably drive the differential types of selection) should differ so drastically between observed and unobserved skills.” (p. 34).²² This argument implies that the prevailing pattern of positive selection on education that is observed in the data is likely to be matched by positive selection on unobservables. Intuitively, the quality-maximizing degree of selectivity (*i*) decreases with the strength of the positive selection on unobservables, which is related to γ , and (*ii*) increases with the return to education $\mu_0^e - \mu_0^u$ that migrants earn at origin. Notice that the differential between π^u and π^e that maximizes migrants’ quality does not depend on the scale of migration. This implies that a change in immigration policy that is meant to change the scale $n(\pi^e, \pi^u)$ of incoming migration flows should always keep $\pi^u - \pi^e = \frac{1-\gamma}{\gamma}(\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$. Recall that the migration rates of educated and uneducated individuals stand at $1 - \Phi(z^e)$ and $1 - \Phi(z^u)$ respectively, so that the quality-maximizing share of educated migrants, which we denote by s_n^e , does vary with n . Specifically, it is straightforward to see that:

$$\frac{\partial s_n^e}{\partial n} = (1 - s_n^e)s_n^e \left(\lambda[g(n)] - \lambda \left[g(n) - \frac{\mu_1^e - \mu_1^u}{\gamma} \right] \right) \frac{\partial g(n)}{\partial n} < 0, \quad (11)$$

where $g(n)$ gives the (unique) value of z^u that determines a scale of migration equal to n when z^e is at its quality-maximizing value.²³ Hence, (11) implies that the share of educated migrants that maximizes migrants’ quality is a *decreasing* function of the scale of migration n : a destination country wishing to increase the scale of incoming migration flows should let the number of uneducated migrants grow proportionally more than the number of educated migrants. To put it differently, expanding the scale of migration exclusively (or mostly) through an expansion of the number of educated migrants does not maximize migrants’ quality.

The perverse effect demonstrated by Proposition 1 arises because the scale-invariant increase in selectivity reduces the conditional expected value of $\ln(w_{i1}^e)$, while it increases the conditional expected value of $\ln(w_{i1}^u)$, as implied by (7). Nevertheless, migrants’ quality begins to decline with a scale-invariant increase in selectivity already when educated migrants earn, on average, more than uneducated migrants, as demonstrated by the following

²²For instance, the analysis conducted by Fernández-Huertas Moraga (2011) uncovers a similar pattern of (negative) selection on observables and unobservables of Mexican migrants to the United States.

²³Notice that $\partial g(n)/\partial n < 0$; this, together with the fact that $\lambda(\cdot)$ is a monotonically increasing function, suffices to sign (11).

Corollary to Proposition 1:

Corollary 2 *The expected value of the log wage at destination is higher for educated than for uneducated migrants when $\pi^u - \pi^e = \frac{1-\gamma}{\gamma}(\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$.*

Proof. See Appendix A.2.

Corollary 2 demonstrates that an increase in selectivity can backfire even though educated agents still have a higher wage than uneducated migrants.

No perverse effects of selectivity can arise when migrants are negatively selected on unobservables, as demonstrated by the following Corollary:

Corollary 3 *If migrants' are negatively selected on unobservables, then a scale-preserving increase in selectivity always increases migrants' quality.*

Proof. See Appendix A.3.

A scale-preserving increase in selectivity unambiguously increases migrants' quality when migrants are negatively selected on unobservables, as it increases the differential in the conditional expected value of the log wage for educated and uneducated migrants, an effect that reinforces the direct positive effect of selectivity on observable characteristics on migrants' quality. This, in turn, implies that the empirical relevance of our theoretical prediction that more selective immigration policies can reduce migrants' quality rests on the hypothesis that migrants are positively selected on unobservables.

4 Extensions

We consider here four extensions of the basic specification of our model that do not influence the theoretical prediction that a scale-invariant increase in selectivity can actually reduce migrants' quality. Specifically, we discuss the implications of (i) introducing general equilibrium effects, (ii) introducing unobserved heterogeneity in time-equivalent migration costs, (iii) allowing for a greater role of uncertainty in the location-decision problem that agents face, and (iv) considering a change in the informational structure for educated agents.

4.1 General equilibrium effects

Our analysis assumes, as in Borjas (1987, 1999b), that the returns to observable and unobservable characteristics in the two countries do *not* respond to migration flows, i.e., the distributions of $\ln(w_0^l)$ and $\ln(w_1^l)$, with $l \in \{e, u\}$, are exogenous.²⁴ If we allowed wages at destination to react to the scale of migration of educated and uneducated agents, then this would strengthen the theoretical prediction contained in Proposition 1, as a scale-invariant increase in selectivity would, through general equilibrium effects, reduce the difference in the expected value of the two conditional distributions.²⁵

4.2 Random variation in time-equivalent migration costs

We retained the assumption that time-equivalent migration costs do not vary across agents with the same observable characteristics, so that self-selection into migration is based only on (observed and unobserved) factors that influence the wages in the two countries. Still, people “are often genuinely reluctant to leave familiar surrounding” (Sjaastad, 1962, p. 85) and “also move for noneconomic reasons” (Chiswick, 1999, p. 184), and this calls for extending the model with the inclusion of heterogeneity in the preferences for migration. We can follow Borjas (1999b) by assuming that time-equivalent migration costs π_i^l are determined by the realization of a normal random variable, i.e., $\pi_i^l = \mu_\pi^l + \epsilon_{i\pi}$, possibly correlated with ϵ_{i0} and ϵ_{i1} . This extension implies that the probability to migrate is given by:

$$\Pr(\tilde{\epsilon}_{i2} \equiv \epsilon_{i1} - \epsilon_{i0} - \epsilon_{i\pi} > \mu_0^l + \mu_\pi^l - \mu_1^l) = 1 - \Phi(\tilde{z}^l),$$

where $\tilde{z}^l = [\mu_0^l + E(\pi^l) - \mu_1^l] / \tilde{\sigma}_2$ and $\tilde{\sigma}_2 = (\sigma_1^2 + \sigma_0^2 + \sigma_\epsilon^2 - 2\sigma_{01} + 2\sigma_{0\pi} - 2\sigma_{1\pi})^{1/2}$. Notice that if the unobserved heterogeneity in the preferences for migration is uncorrelated with the unobservables that influence wages, i.e., $\sigma_{0\pi} = \sigma_{1\pi} = 0$,²⁶ then $|\tilde{\gamma}| \equiv (\sigma_1^2 - \sigma_{01}) / \tilde{\sigma}_2 < |\gamma|$, as $\tilde{\sigma}_2 > \sigma_2$. The differential between the conditional expectation of $\ln(w_1^l)$ and μ_1^l is equal to:

²⁴“We do not yet understand the nature of selection in a general equilibrium context.” (Borjas, 2014, p. 213).

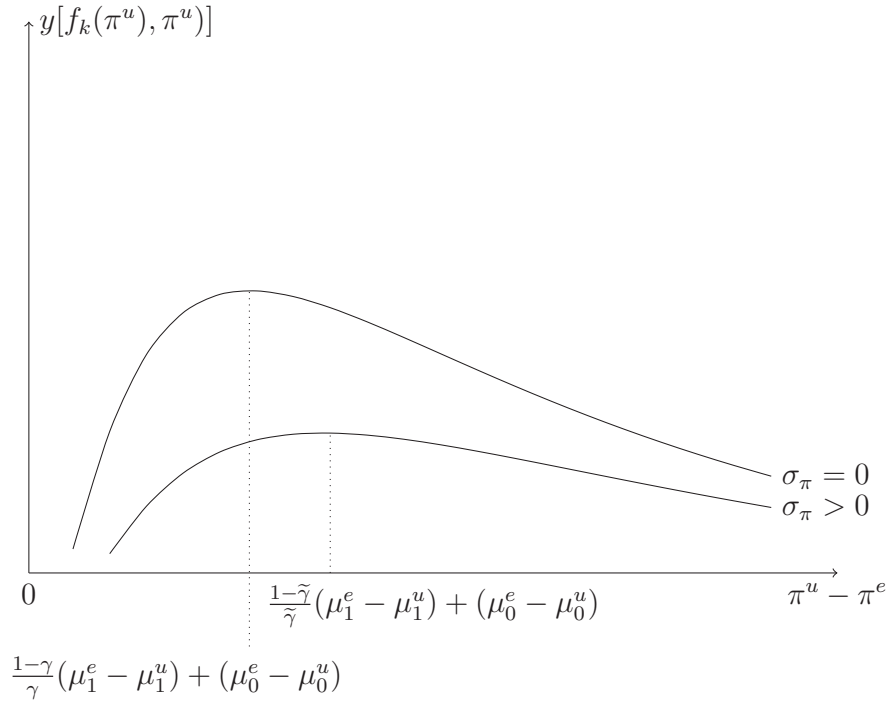
²⁵This reasoning implicitly rests on the assumption of complementarity between educated and uneducated migrants in the labor market at destination. Note that allowing for general equilibrium effects renders problematic our reliance on the logarithm of migrants’ wages as a measure of their quality.

²⁶“[T]he nature of the correlation between costs (whether in absolute or in time-equivalent terms) and skills is unclear.” (Borjas, 2014, p. 10).

$$\tilde{Q}_1^l = \left(\tilde{\gamma} - \frac{\sigma_{1\pi}}{\tilde{\sigma}_2} \right) \lambda(\tilde{z}^l). \quad (12)$$

In the absence of covariance between the stochastic component of the log wage at destination and the stochastic component of migration costs, then we obtain the same pattern of selection on unobservables than in Section 2, but with $|\tilde{Q}_1^l| < |Q_1^l|$ as $|\tilde{\gamma}| < |\gamma|$ and $\tilde{z}^l < z^l$. Intuitively, self-selection on noneconomic factors (preferences for migration), dilutes the extent of self-selection on unobserved ability. This, in turn, increases the scope for quality-enhancing scale-preserving increases in selectivity when migrants are positively self-selected on unobservables, i.e., $\tilde{Q}_1^l > 0$, as the indirect adverse effect of the policy change becomes weaker, as depicted in Figure 3.²⁷

Figure 3: Unobserved heterogeneity in preferences for migration



Note: the figure represents the evolution of migrants' quality for a given scale of migration for two different values of σ_π , with $\sigma_{0\pi} = \sigma_{1\pi} = 0$ and $\gamma > 0$.

The literature suggests that there is a negative correlation between migration costs and

²⁷We have that the quality-maximizing share of educated migrants s_n^e is an increasing function of σ_π .

wages (see, for instance, Chiswick (1999), Bellettini and Berti Ceroni (2007), Chiquiar and Hanson (2005), McKenzie and Rapoport (2010) and Beine *et al.* (2011)) and this would widen the scope for a positive selection on unobservables of the migrants. The term between brackets in (12), which determines the pattern of migrants' self-selection on unobservables, can be positive even if $\tilde{\gamma} \leq 0$, when the correlation between time-equivalent migration costs π^l and ϵ_1 is negative. Since a pattern of positive selection on unobservables represents a necessary condition to obtain Proposition 1, introducing heterogeneity in preferences for migration would strengthen our theoretical prediction under the empirically relevant assumption that $\rho_{1\pi} < 0$. The optimal differential between z^e and z^u can be either above or below the value that corresponds to the absence of heterogeneity in the preferences for migration.

4.3 An alternative informational structure

As in Borjas (1987, 1999b), we have assumed, that wages are *remotely* observable, so that the information set upon which the decision to migrate is taken includes the realizations of both ϵ_{i0} and ϵ_{i1} . Bertoli (2010b) considers an alternative informational structure where only ϵ_{i0} belongs to the information set of the agents while the realization of ϵ_{i1} is not observed before migrating. Agents are assumed to know the parameters that characterized the bivariate normal distribution of $\ln(w_{i0}^l)$ and $\ln(w_{i1}^l)$, so that the realization of ϵ_{i0} conveys, in general, information on the expected value of the stochastic component of $\ln(w_{i1}^l)$. With this alternative informational structure, the probability of migrating is given by:^{28,29}

$$\Pr \left[\left(\frac{\sigma_1}{\sigma_0} \rho_{01} - 1 \right) \epsilon_{i0} > \mu_0^l + \pi^l - \mu_1^l \right] = \begin{cases} 1 - \Phi(\hat{z}^l) & \text{if } \rho_{01} > \sigma_0/\sigma_1 \\ \Phi(\hat{z}^l) & \text{if } \rho_{01} < \sigma_0/\sigma_1 \end{cases}$$

where:

²⁸Migrants have domestic wages belonging to the lower (upper) tail of the truncation of $\ln(w_{i0}^l)$ when $\rho_{01} > \sigma_0/\sigma_1$ ($\rho_{01} < \sigma_0/\sigma_1$). When $\rho_{01} = \sigma_0/\sigma_1$, there is no heterogeneity in migration decisions, i.e., all agents stay (migrate) if $\hat{z} > 0$ ($\hat{z} < 0$), as the realization of ϵ_{i0} coincides with the conditional expectation of ϵ_{i1} .

²⁹The probability of self-selection into migration is always lower when wages are only locally rather than remotely observable if less than half of the population at origin migrates, i.e., $\mu_0 + \pi - \mu_1 > 0$. This follows from the fact that $[(\sigma_1 \rho_{01} / \sigma_0) - 1] \sigma_0 < \sigma_2$ whenever $\rho_{01} \in (-1, 1)$.

$$\hat{z}^l = \frac{\mu_0^l + \pi^l - \mu_1^l}{\sigma_0 \left(\frac{\sigma_1}{\sigma_0} \rho_{01} - 1 \right)}$$

Bertoli (2010, p. 91) demonstrates that:

$$\hat{Q}_1^l = \hat{\gamma} \lambda(\hat{z}^l)$$

where:

$$\hat{\gamma} = \begin{cases} \frac{\sigma_{01}}{\sigma_0} & \text{if } \rho_{01} > \sigma_0/\sigma_1 \\ -\frac{\sigma_{01}}{\sigma_0} \frac{1-\Phi(\hat{z}^l)}{\Phi(\hat{z}^l)} & \text{if } \rho_{01} < \sigma_0/\sigma_1 \end{cases}$$

Thus, when wages are only locally observable, migrants are positively selected on unobservables if and only if $\rho_{01} > \sigma_0/\sigma_1$ or $\rho_{01} < 0$. The alternative informational structure adopted by Bertoli (2010b) reduces the scope for a positive selection on unobservables compared to Borjas (1987), as depicted in Figure 4, but it does not affect the result in Proposition 1: a scale-invariant increase in selectivity can reduce migrants' quality when migrants' are positively selected on unobservables. Specifically, when wages are only locally observable and positively selected on unobservables, then a scale-preserving increase in selectivity reduces migrants' quality if and only if:

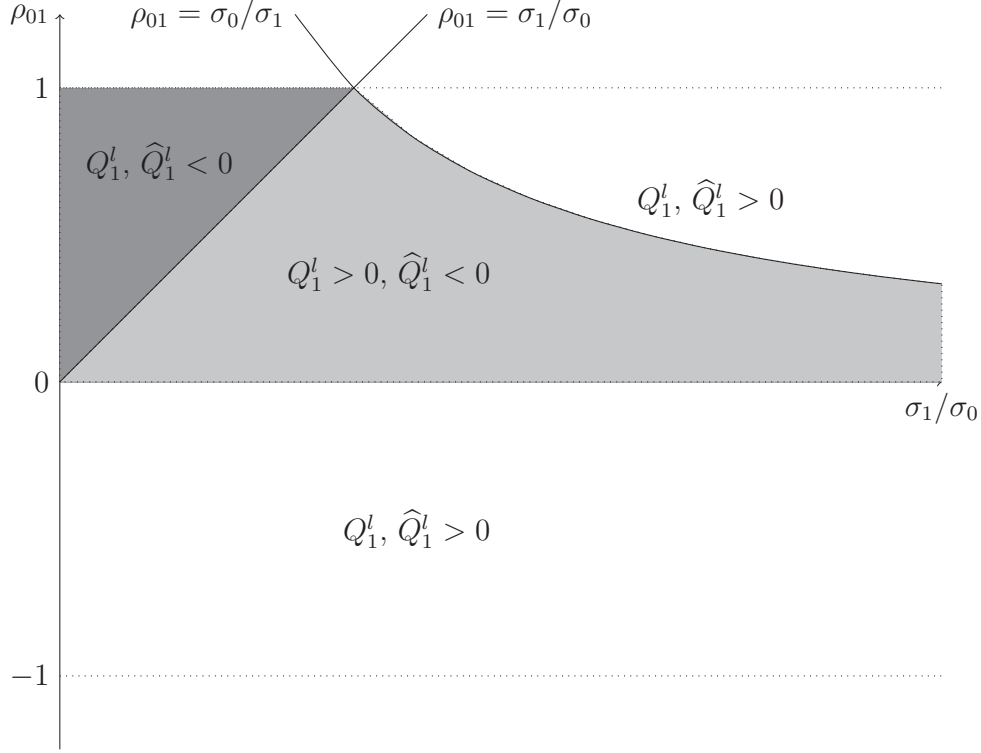
$$\pi^u - \pi^e > \frac{1 - \hat{\gamma}}{\hat{\gamma}} (\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$$

As discussed in the next section, the sign of the difference between γ and $\hat{\gamma}$ depends, in general, on the elements of the covariance matrix Σ and on the scale of migration, so a change in the informational structure influences in an ambiguous way the quality-maximizing differential $\pi^u - \pi^e$.

4.4 Educated migrants arriving “with a job in hand”

Borjas and Friedberg (2009) suggest that high-skilled immigrants who enter into the United States with a H1-B visa have a higher quality (initial relative wage) as “arriving with a job in hand eliminates some of the initial labor market disadvantage of new immigrants” (p. 21), and this contributes to explain the observed uptick in immigrants' quality in 2000. Selective policies could act not only on the cost side, as we have assumed so far in our analysis, but also on the size of the information set upon which the decision to migrate is taken. Such

Figure 4: Selection on unobservables in Borjas (1987) and Bertoli (2010b)



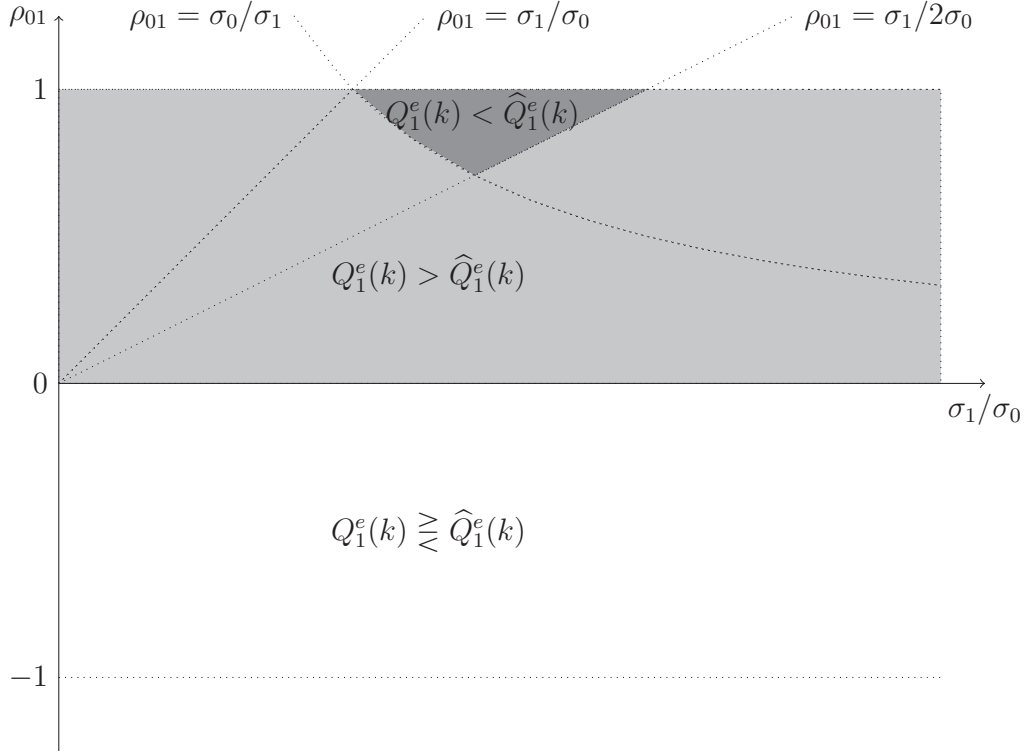
a change in the informational structure has an influence on both the scale of migration and on migrants' selection on unobservables. We can analyze its effects by assuming that the informational structure changes from that of Bertoli (2010b) to that of Borjas (1987),³⁰ so that wages become *remotely* observable for educated potential migrants, who can arrive “with a job in hand”. We can also assume that the destination country adjusts migration costs for educated migrants in order to keep the scale of migration unchanged, so that the change in the informational structure is scale-preserving.

We have that better information *reduces* migrants' quality when $\rho_{01} > \sigma_0/\sigma_1$ and $\rho_{01} > \sigma_1/2\sigma_0$ or when $\rho_{01} < 0$ if the scale k of migration is sufficiently small,³¹ as depicted in Figure 5. Remarkably, the proposed change in the informational structure is detrimental for migrants' quality when unobservable skills can be easily transferred across countries, i.e., ρ_{01}

³⁰Specifically, migration costs π^e have to be increased to keep the scale of migration unchanged when $\Phi(z^e) > 1/2$ and $\rho_{01} \in (-1, 1)$.

³¹See Appendix A.4 for a derivation of these results.

Figure 5: Change in the informational structure and migrants' quality



Note: the figure is drawn under the assumption that $\mu_0^e + \pi^e > \mu_1^e$.

is high, and the destination country offers a reward to ability that can be up to twice as large as the one at origin. Hence, expanding the policy instruments that destination countries have at their disposal can either weaken or strengthen our argument that an increase in selectivity can reduce migrants' quality.

Notice also that once we introduce different informational structures for educated and uneducated agents,³² then we have that different patterns of selection on unobservables for the two groups can prevail. Specifically, as implied by Figure 4, educated migrants are positively selected while uneducated migrants are negatively selected on unobservables when $0 < \rho_{01} < \min\{\sigma_1/\sigma_0, \sigma_0/\sigma_1\}$. When this is the case, then a scale-preserving increase in selectivity reduces the average wage for both educated and uneducated migrants. This

³²Specifically, wages are remotely observable *à la* Borjas (1987) for educated agents and only locally observable *à la* Bertoli (2010b) for uneducated agents

occurs as the reduction in migration costs for educated agents dilutes their positive selection on unobservables, while the simultaneous increase in migration costs for uneducated agents exacerbates the intensity of their negative selection on unobservables. This, in turn, strengthens our theoretical prediction the possible detrimental impact on migrants' quality of an increase in selectivity.

5 Conclusion

The effect on migrants' quality produced by an increase in the selectivity of immigration policies based on potential migrants' observable characteristics crucially depends on how the policy change influences migrants' selection on unobservables, such as ability and motivation, that influence their wages at destination. Our theoretical model shows that a scale-preserving increase in the share of educated migrants can actually reduce migrants' quality when migrants have, on average, a higher level of ability than stayers. The relevance of individual characteristics that remain *unobserved* for immigration officers in explaining observed differences in earnings suggest that the scope for perverse effects of selective immigration policies could be more than a theoretical curiosity.

References

- AMBROSINI, J. W., AND G. PERI (2012): "The Determinants and the Selection of Mexico-US Migrants," *World Economy*, 35(2), 111–151.
- ANTECOL, H., D. A. COBB-CLARK, AND S. J. TREJO (2003): "Immigration Policy and the Skills of Immigrants to Australia, Canada, and the United States," *Journal of Human Resources*, XXXVIII(1), 192–218.
- AYDEMIR, A. (2011): "Immigrant selection and short-term labor market outcomes by visa category," *Journal of Population Economics*, 24(2), 451–475.
- AYDEMIR, A., AND M. SKUTERUD (2005): "Explaining the Deteriorating Entry Earnings of Canada's Immigrant Cohorts, 1966 - 2000," *Canadian Journal of Economics*, 38(2), 641–671.

- BEINE, M., F. DOCQUIER, AND Ç. ÖZDEN (2011): “Diasporas,” *Journal of Development Economics*, 95(1), 30–41.
- BEINE, M., F. DOCQUIER, AND H. RAPOPORT (2001): “Brain Drain and Economic Growth: Theory and Evidence,” *Journal of Development Economics*, 64, 275–289.
- BELLETTINI, G., AND C. BERTI CERONI (2007): “Immigration Policy, Self-selection, and the Quality of Immigrants,” *Review of International Economics*, 15(5), 869–877.
- BELLOT, M., AND T. HATTON (2012): “Skill Selection and Immigration in OECD Countries,” *Scandinavian Journal of Economics*, 114(4), 1105–1128.
- BERTOLI, S. (2010a): “Networks, sorting and self-selection of Ecuadorian migrants,” *Annales d’Économie et de Statistique*, 97/98, 261–288.
- (2010b): “The Informational Structure of Migration Decision and Migrants’ Self-Selection,” *Economics Letters*, 108(1), 89–92.
- BERTOLI, S., AND H. BRÜCKER (2011): “Selective Immigration Policies, Migrants’ Education and Welfare at Origin,” *Economics Letters*, 113(1), 19–22.
- BERTOLI, S., H. BRÜCKER, G. FACCHINI, A. MAYDA, AND G. PERI (2012): “Understanding Highly Skilled Migration in Developed Countries: The Upcoming Battle for Brains,” in *Brain Drain and Brain Gain: The Global Competition to Attract High-Skilled Migrants*, ed. by T. Boeri, H. Brücker, F. Docquier, and H. Rapoport, pp. 15–198. Oxford University Press.
- BERTOLI, S., J. FERNÁNDEZ-HUERTAS MORAGA, AND F. ORTEGA (2013): “Crossing the Border: Self-Selection, Earnings and Individual Migration Decisions,” *Journal of Development Economics*, 101, 75–91.
- BERTOLI, S., AND H. RAPOPORT (2015): “Heaven’s Swing Door: Endogenous Skills, Migration Networks and the Effectiveness of Quality-Selective Immigration Policies,” *Scandinavian Journal of Economics*, forthcoming.
- BIANCHI, M. (2013): “Immigration Policy and Self-Selecting Migrants,” *Journal of Public Economic Theory*, 15(1), 1–23.

- BIAVASCHI, C., AND B. ELSNER (2013): “Let’s Be Selective about Migrant Self-Selection,” IZA Discussion Paper No. 7865, Bonn.
- BORJAS, G. J. (1985): “Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants,” *Journal of Labor Economics*, 3(4), 463–489.
- (1987): “Self-Selection and the Earnings of Immigrants,” *American Economic Review*, 77(4), 531–553.
- (1999a): *Heaven’s Door: Immigration Policy and the American Economy*. Princeton: Princeton University Press.
- (1999b): “The Economic Analysis of Immigration,” in *Handbook of Labor Economics, Volume 3A*, ed. by O. C. Ashenfelter, and D. Card, pp. 1697–1760. North-Holland.
- (2014): *Immigration Economics*. Cambridge, MA: Harvard University Press.
- BORJAS, G. J., AND R. M. FRIEDBERG (2009): “Recent Trends in the Earnings of New Immigrants to the United States,” NBER Working Paper No. 15406.
- CARD, D. (1990): “The Impact of the Mariel Boatlift on the Miami Labor Market,” *Industrial and Labor Relations Review*, 43(2), 245–257.
- CHIQUIAR, D., AND G. H. HANSON (2005): “International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States,” *Journal of Political Economy*, 113(2), 239–281.
- CHISWICK, B. R. (1999): “Are Immigrants Favorably Self-Selected?,” *American Economic Review*, 89(2), 181–185.
- DEQUIEDT, V., AND Y. ZENOU (2013): “International Migration, Imperfect Information, and Brain Drain,” *Journal of Development Economics*, 102, 62–78.
- DOCQUIER, F., O. FAYE, AND P. PESTIEAU (2008): “Is Migration a Good Substitute for Education Subsidies?,” *Journal of Development Economics*, 86, 263–276.
- DOCQUIER, F., B. L. LOWELL, AND A. MARFOUK (2009): “A Gendered Assessment of Highly Skilled Emigration,” *Population and Development Review*, 35(2), 297–321.

- DOCQUIER, F., J. MACHADO, AND K. SEKKAT (2015): “Efficiency Gains from Liberalizing Labor Mobility,” *Scandinavian Journal of Economics*, forthcoming.
- FERNÁNDEZ-HUERTAS MORAGA, J. (2011): “New Evidence on Immigrant Selection,” *Review of Economics and Statistics*, 93(1), 72–96.
- (2013): “Understanding Different Migrant Selection Patterns in Rural and Urban Mexico,” *Journal of Development Economics*, 103, 182 – 201.
- GIORDANI, P. E., AND M. RUTA (2013): “Coordination Failures in Immigration Policy,” *Journal of International Economics*, 89(1), 55 – 67.
- GROGGER, J., AND G. H. HANSON (2011): “Income Maximization and the Selection and Sorting of International Migrants,” *Journal of Development Economics*, 95, 42–57.
- HECKMAN, J. J. (1979): “Sample Selection Bias as a Specification Error,” *Econometrica*, 47(1), 153–161.
- HOLMSTROM, B. AND P. MILGROM (1991): “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design,” *Journal of Law, Economics and Organization*, 7, 24–52.
- JASSO, G., AND M. R. ROSENZWEIG (2009): “Selection Criteria and the Skill Composition of Immigrants: A Comparative Analysis of Australian and U.S. Employment Immigration,” in *Skilled Migration Today: Phenomenon, Prospects, Problems, Policies*, ed. by J. Bhagwati, and G. Hanson, pp. 153–183. Oxford University Press.
- KAESTNER, R., AND O. MALAMUD (2014): “Self-Selection and International Migration: New Evidence from Mexico,” *Review of Economics and Statistics*, 96(1), 78–91.
- MCKENZIE, D. J. (2007): “Paper Walls Are Easier to Tear Down: Passport Costs and Legal Barriers to Emigration,” *World Development*, 35(11), 2026 – 2039.
- MCKENZIE, D. J., AND H. RAPOPORT (2010): “Self-selection patterns in Mexico-U.S. migration: The role of migration networks,” *The Review of Economics and Statistics*, 92(4), 811–821.

- MOUNTFORD, A. (1997): "Can a brain drain be good for growth in the source economy?," *Journal of Development Economics*, 53(2), 287–303.
- ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.
- SCHULTZ, T. W. (1975): "The Value of the Ability to Deal with Disequilibria," *Journal of Economic Literature*, 13(3), 827–846.
- SJAASTAD, L. A. (1962): "The Costs and Returns of Human Migration," *Journal of Political Economy*, 70(1), 80–93.

A Appendix

A.1 Proof of Proposition 1

We can rely on the definition of the Inverse Mills ratio $\lambda(z^l)$, with $l \in \{e, u\}$, to rewrite (10) as follows:

$$y(\pi^e, \pi^u) = \frac{1}{n(\pi^e, \pi^u)} [\alpha ([1 - \Phi(z^e)]\mu_1^e + \gamma\phi(z^e)) + (1 - \alpha) ([1 - \Phi(z^u)]\mu_1^u + \gamma\phi(z^u))]$$

Deriving this expression with respect to π^u , with $\pi^e = f_k(\pi^u)$, and using the fact that $\partial\phi(z^l)/\partial z^l = -z^l\phi(z^l)$, we get:

$$\frac{\partial y[f_k(\pi^u), \pi^u]}{\partial \pi^u} = \frac{(1 - \alpha)\phi(z^u)}{\sigma_2 n(\pi^e, \pi^u)} [\mu_1^e - \mu_1^u + \gamma(z^e - z^u)]$$

The fraction on the right hand side of the expression above is always positive, so that its sign is determined by the sign of the term between brackets. When migrants are positively selected on unobservables, i.e., $\gamma > 0$, the term between brackets is negative if and only if

$$z^e - z^u < -\frac{\mu_1^e - \mu_1^u}{\gamma}$$

Using the definitions of z^e and z^u , the inequality can be rewritten as:

$$\pi^u - \pi^e > \frac{1 - \gamma}{\gamma} (\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$$

This concludes the proof. ■

A.2 Proof of Corollary 2

Recall that $\pi^u - \pi^e = \frac{1 - \gamma}{\gamma} (\mu_1^e - \mu_1^u) + (\mu_0^e - \mu_0^u)$ implies that:

$$z^e = z^u - \frac{\mu_1^e - \mu_1^u}{\gamma}$$

We want to show that:

$$E \left[\ln(w_{i1}^e) | \epsilon_{i2}^e > z^u - \frac{\mu_1^e - \mu_1^u}{\gamma} \right] > E \left[\ln(w_{i1}^u) | \epsilon_{i2}^u > z^u \right]$$

From (4), we can rewrite the inequality above as follows:

$$\mu_1^e + \gamma \lambda \left(z^u - \frac{\mu_1^e - \mu_1^u}{\gamma} \right) > \mu_1^u + \gamma \lambda(z^u)$$

Moving terms around and remembering that Proposition 1 rests on the assumption that $\gamma > 0$, we obtain:

$$\lambda(z^u) - \lambda \left(z^u - \frac{\mu_1^e - \mu_1^u}{\gamma} \right) < -\frac{\mu_1^e - \mu_1^u}{\gamma}$$

The inequality above holds as the Inverse Mills ratio $\lambda(z^u)$ is a contraction mapping. ■

A.3 Proof of Corollary 3

Let us go back to the partial derivative of migrants' quality with respect to π^u , with $\pi^e = f_k(\pi^u)$:

$$\frac{\partial y[f_k(\pi^u), \pi^u]}{\partial \pi^u} = \frac{(1 - \alpha)\phi(z^u)}{\sigma_2 n(\pi^e, \pi^u)} [\mu_1^e - \mu_1^u + \gamma(z^e - z^u)]$$

When $\gamma < 0$, the term between brackets is positive if:

$$z^e - z^u < -\frac{\mu_1^e - \mu_1^u}{\gamma}$$

As $z^e < z^u$ and the right hand side of the inequality above is positive, then this condition is always satisfied. ■

A.4 Change in the informational structure and migrants' quality

A scale-preserving change in the informational structure, with wages for educated individuals only being not just locally but also remotely observable, increases educated migrants' quality, i.e., $Q_1^e(k) > \widehat{Q}_1^e(k)$, if and only if:

$$\frac{\sigma_{12}}{\sigma_2} = \gamma > \widehat{\gamma} = \begin{cases} \frac{\sigma_{01}}{\sigma_0} & \text{if } \rho_{01} > \sigma_0/\sigma_1 \\ -\frac{1-\Phi(\widehat{z})}{\Phi(\widehat{z})} \frac{\sigma_{01}}{\sigma_0} & \text{if } \rho_{01} < \sigma_0/\sigma_1 \end{cases} \quad (13)$$

where \widehat{z}^e gives rise to a scale of migration equal to k under the informational structure in ?. We have that (13) clearly holds when $\widehat{\gamma} < 0 < \gamma$, i.e., $\rho_{01} < \min\{\sigma_1/\sigma_0, \sigma_0/\sigma_1\}$. When $\rho_{01} > \sigma_0/\sigma_1$, then (13) can be rewritten as:

$$\frac{\sigma_1^2 - \sigma_{01}}{(\sigma_1^2 + \sigma_0^2 - 2\sigma_{01})^{1/2}} > \frac{\sigma_{01}}{\sigma_0}$$

Moving terms around, and taking both sides to the power of two, we obtain:

$$\sigma_0^2 \sigma_1^2 (\sigma_1^2 - 2\sigma_{01}) > \sigma_{01}^2 (\sigma_1^2 - 2\sigma_{01}) \quad (14)$$

If $\rho_{01} < \sigma_1/2\sigma_0$, then (14) is equivalent to:

$$\sigma_0^2 \sigma_1^2 > \sigma_{01}^2$$

which clearly holds as long as $\rho_{01} < 1$. If $\rho_{01} > \sigma_1/2\sigma_0$, then (14) simplifies to:

$$\sigma_0^2 \sigma_1^2 < \sigma_{01}^2$$

which cannot hold. Hence, when $\rho_{01} > \sigma_0/\sigma_1$, we have that $Q_1^e(k) > \widehat{Q}_1^e(k)$ when $\rho_{01} < \sigma_1/2\sigma_0$, while $Q_1(k) < \widehat{Q}_1^e(k)$ when $\rho_{01} > \sigma_1/2\sigma_0$.

When $\rho_{01} < 0$, we can demonstrate that the sign of the difference between $Q_1^e(k)$ and $\widehat{Q}_1^e(k)$ is ambiguous, and dependent on the scale of migration k , and hence implicitly on \widehat{z}^e . Specifically, following the previous steps, we can show that:

$$\frac{\sigma_{12}}{\sigma_2} > -\frac{\sigma_{01}}{\sigma_0}$$

but this does not allow to sign:

$$\frac{\sigma_{12}}{\sigma_2} \begin{matrix} \geq \\ < \end{matrix} -\frac{1 - \Phi(\widehat{z}^e)}{\Phi(\widehat{z}^e)} \frac{\sigma_{01}}{\sigma_0} \quad (15)$$

unless we introduce assumptions on the value of \widehat{z}^e . Specifically, (15) implicitly defines a threshold, which is always positive and that depends on the elements of the covariance matrix Σ , such that $Q_1^e(k)$ is higher (lower) than $\widehat{Q}_1^e(k)$ when \widehat{z}^e is below (above) this threshold.

Finally, when $\rho_{01} > \sigma_1/\sigma_0$, we can demonstrate that:

$$\frac{\sigma_{12}}{\sigma_2} > -\frac{\sigma_{01}}{\sigma_0} > -\frac{1 - \Phi(\widehat{z}^e)}{\Phi(\widehat{z}^e)} \frac{\sigma_{01}}{\sigma_0} \quad (16)$$

when $\widehat{z}^e > 0$. Again, we have that the first inequality in (16) is satisfied if and only if:

$$-\frac{\sigma_{12}}{\sigma_2} < \frac{\sigma_{01}}{\sigma_0}$$

Moving terms around, and taking both sides to the power of two, we obtain:

$$\sigma_0^2 \sigma_1^2 (\sigma_1^2 - 2\sigma_{01}) < \sigma_{01}^2 (\sigma_1^2 - 2\sigma_{01}) \quad (17)$$

As $\rho_{01} > \sigma_1/2\sigma_0$, then (17) is equivalent to:

$$\sigma_0^2 \sigma_1^2 > \sigma_{01}^2$$

which clearly holds. Hence, $Q_1^e(k) > \widehat{Q}_1^e(k)$ when $\rho_{01} > \sigma_1/\sigma_0$, as depicted in Figure 5.