DISCUSSION PAPER SERIES

IZA DP No. 9538

Can Selective Immigration Policies Reduce Migrants' Quality?

Simone Bertoli
Vianney Dequiedt
Yves Zenou

November 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Can Selective Immigration Policies Reduce Migrants' Quality?

Simone Bertoli
CERDI, University of Auvergne, CNRS and IZA

Vianney Dequiedt
CERDI, University of Auvergne

Yves Zenou
Stockholm University, IFN and IZA

Discussion Paper No. 9538
November 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 9538
November 2015

ABSTRACT

Can Selective Immigration Policies Reduce Migrants' Quality?*

Destination countries can adopt selective immigration policies to improve migrants' quality. Screening potential migrants on the basis of observable characteristics also influences their self-selection on unobservables. We propose a model that analyzes the effects of selective immigration policies on migrants' quality, measured by their wages at destination. 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 on unobservables. We also demonstrate that, in this case, the quality-maximizing share of educated migrants declines with the scale of migration.

JEL Classification: F22, K37, J61

Keywords: selective policies, self-selection, migrants' quality

Corresponding author:

Simone Bertoli
CERDI
University of Auvergne
Bd. F. Mitterrand, 65
F-63000 Clermont-Ferrand
France
E-mail: simone.bertoli@udamail.fr

* The authors are grateful to the editor Kaivan Munshi, an anonymous referee, Jesús Fernández-Huertas Moraga, Hillel Rapoport and the participants at various conferences and seminars for their helpful comments and suggestions. Simone Bertoli and Vianney Dequiedt acknowledge the support received from the FERDI and the *Agence Nationale de la Recherche* of the French government through the program "*Investissements d'avenir*" (ANR-10-LABX-14-01). The usual disclaimers apply.

"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* (p. 215).

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 decades[1] 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, granting better chances of admission at destination to applicants endowed with more desirable individual characteristics.[3]

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, motivation or soft skills (Heckman and Kautz, 2012), remain *un*observable 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

[1] See, for instance, Borjas (1985, 2015) 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).

[3] "The main policy proposals on the agenda are increasing attempts to create a more attractive and favorable regime for highly skilled (or just plain wealthy) migrants." (Pritchett, 2006, pp. 106-107).

may not necessarily "transfer over" to a more comprehensive measure of a worker's human capital" (Borjas, 2014, pp. 29-30).[4] 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).[5,6] 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 analyzes 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, based on Borjas (1987), 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,[7] which is defined as a reduction of migration costs for educated applicants, matched by a simultaneous increase in migration 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

[4] Along the same lines, Kaestner and Malamud (2014, p. 89) caution about the limits of "using individual components of skill such as education to assess migrant selection with respect to skill".

[5] Antecol *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.

[6] Ambrosini 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.

[7] This 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; keeping the scale of migration constant allows not to blur the effects due to a variation in selectivity with the effects produced by a change in the openness of immigration policies.

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 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. We also demonstrate that the share of educated agents among the migrants that maximizes quality is negatively related to the scale of migration when migrants are positively selected on unobservables. If the share of educated migrants is kept unchanged while expanding the scale of migration, then the set of educated agents that are induced to migrate by the reduction in migration costs has a *lower* quality than the corresponding set of uneducated agents, and this difference in quality at the margin is inconsistent with the maximization of migrants' quality.

These theoretical results are robust with respect to several extensions of the basic version of the model. Specifically, we analyze the implications of (*i*) allowing for a greater dispersion in the quality of educated agents, (*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 the migration-decision problem for 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 unfolds 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 discusses its relationship with the empirical literature. Then, Section 4 describes the robustness of our theoretical predictions with respect to various extensions of the model. Finally, Section 5 concludes.

2 The model

We develop a random utility maximization model to describe the location-decision problem that potential migrants face. We consider an origin country, which is denoted by 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 α 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 it influences the agents' wage. Individuals are heterogeneous in other characteristics that also influence their wages, which are exogenous with respect to migration. Specifically, we assume that:

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

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

$$\left(\ln w_{i0}^l, \ln w_{i1}^l\right)' \sim \mathcal{N}(\boldsymbol{\mu}^l, \boldsymbol{\Sigma}^l). \tag{1}$$

We also assume that $\mu_j^e > \mu_j^u$ for $j = 0, 1$, and $\boldsymbol{\Sigma}^e = \boldsymbol{\Sigma}^u$.[8] The wage equation above implies that individual earnings in both countries and for both education levels can be decomposed

[8]The assumption of identical covariance matrices for educated and uneducated agents is relaxed in Section 4.1 below.

into a part due to *observable* characteristics (μ_j^l) and a part due to *unobserved* characteristics (ϵ_{ij}). For the individual 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. 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. 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.[9]

Wages are *remotely* observable, so that agents decide whether to migrate or not after having observed the realizations of the stochastic component of domestic and foreign wages.[10] Migration represents a utility-maximizing decision if and only if:[11]

$$\ln w_{i0}^l + \pi^l \leq \ln w_{i1}^l,$$

where $\pi^l = \ln(1 + C_i/w_{i0}^l)$.[12] Educated and uneducated agents face different time-equivalent migration costs. The probability that migration represents the utility-maximizing option is given by:

$$\Pr\left(\epsilon_{i2} \equiv \epsilon_{i1} - \epsilon_{i0} \geq \mu_0^l + \pi^l - \mu_1^l\right) = \Phi(-z^l), \tag{2}$$

where $\Phi(.)$ 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 ϵ_2. 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

[9]Unobserved heterogeneity in the preferences for migration is introduced in Section 4.2.

[10]We also consider an alternative informational structure, with *locally* observable wages along the lines of Bertoli (2010b); the implications of this alternative informational structure are analyzed in Sections 4.3 and 4.4 below.

[11]Borjas (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.

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

normality implies that (Heckman, 1979; Borjas, 1987):

$$E\left[\ln w_{i1}^{l}|\epsilon_{i2} \geq z^{l}\right] = \mu_{1}^{l} + Q_{1}(z^{l}), \tag{3}$$

with $Q_1(z^l) \equiv \gamma\lambda(z^l)$, where γ is given by the covariance between the conditioning variable ϵ_{i2} and the stochastic component ϵ_{i1} of $\ln w_{i1}^l$, scaled by the standard deviation of the conditioning variable, i.e., $\gamma = (\sigma_1^2 - \sigma_{01})/\sigma_2$, and where $\lambda(z^l) \equiv \phi(z^l)/\Phi(-z^l)$ represents the Inverse Mills ratio. The Inverse Mills ratio corresponds to the expected value of the upper tail of a truncated standard normal distribution, and it is thus a positive and increasing function of z^l, with $\lambda(z^l) > z^l$. We say that the migrants with a level of education l are positively selected on unobservables if $Q_1(z^l) > 0$, and negatively selected when $Q_1(z^l) < 0$. The pattern of selection on unobservables depends exclusively on γ, while the intensity of selection on unobservables depends on z^l, i.e., on the deterministic components of the log wages μ_0^l and μ_1^l, and on time-equivalent migration costs π^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 instance, Giordani and Ruta (2013), Bianchi (2013) and Docquier *et al.* (2015).[13,14]

Destination countries can impose different migration costs on potential migrants with different observable characteristics, such as education. Variations in education-specific migration costs can influence both the scale of migration, and migrants' quality,[15] as they

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

[14] 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.

[15] We follow the literature by defining quality as the average log wage that migrants earn at destination; see, for instance, Borjas (1985) or Aydemir (2011).

modify the intensity of selection on unobservables of both groups. We first characterize the immigration policy (π^u, π^e),[16] and hence the resulting share of educated individuals among the migrants, that maximizes migrants' quality for a given scale of migration, and we then analyze how the quality-maximizing share of educated migrants varies with the scale of migration.

3.1 Quality-maximizing policy for a given scale of migration

We define migrants' quality as a weighted average of the log wages for the two types of migrants:

$$y(z^u, z^e) \equiv \beta(z^u, z^e)\left[\mu_1^e + Q_1(z^e)\right] + \left[1 - \beta(z^u, z^e)\right]\left[\mu_1^u + Q_1(z^u)\right], \tag{4}$$

where the weights are given by the (endogenous) share of educated migrants that, by the law of large numbers, is given by:

$$\beta(z^u, z^e) = \frac{\alpha\Phi(-z^e)}{\kappa(z^u, z^e)}, \tag{5}$$

with

$$\kappa(z^u, z^e) \equiv \alpha\Phi(-z^e) + (1 - \alpha)\Phi(-z^u) \tag{6}$$

representing the scale of migration flows. Using (6), we can define the family of iso-migration curves as:

$$g_k(z^u) \equiv -\Phi^{-1}\left[\frac{k - \Phi(-z^u)(1 - \alpha)}{\alpha}\right], \tag{7}$$

indexed by k, which represents the scale of migration. The iso-migration curves $z^e = g_k(z^u)$ are downward sloping in the (z^u, z^e) immigration policy space, and higher curves correspond to a smaller scale of migration (see Figure 2 below).

We define a scale-preserving increase in selectivity as an increase in the time-equivalent migration cost π^u for uneducated agents, and hence in z^u, along an iso-migration curve. Intuitively, the share of educated migrants $\beta(z^u, z^e)$ monotonically increases with z^u along an iso-migration curve, as from (5) we have that:

$$\frac{\partial \ln \beta[z^u, g_k(z^u)]}{\partial z^u} = -\frac{\partial g_k(z^u)}{\partial z^u}\lambda[g_k(z^u)] > 0, \tag{8}$$

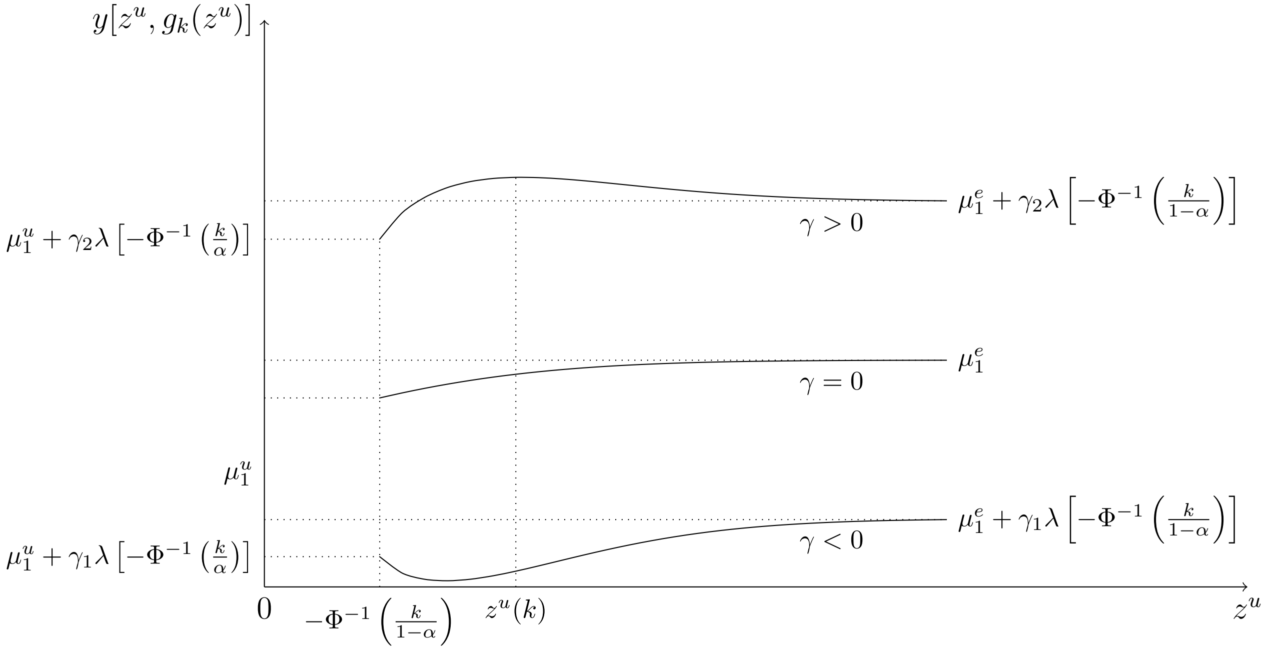
[16] As z^l is a linear function of π^l, the characterization of the optimal immigration policy can be indifferently conducted with respect to (π^u, π^e) or the induced pair (z^u, z^e).

A scale-preserving increase in selectivity influences, in general, migrants' quality in (4) through two distinct channels: (i) it increases the share of educated migrants, whose log wages are drawn from a distribution with a higher unconditional expected value, and (ii) it modifies the intensity of selection for both educated and uneducated migrants. The combined effect of (i) and (ii) is ambiguous whenever $\gamma \neq 0$, as demonstrated by the following Proposition:

Proposition 1 *Migrants' quality is a non-monotonic function of z^u along any iso-migration curve whenever migrants are not randomly selected on unobservables.*

Proof. See Appendix A.1. ∎

Figure 1: Migrants' quality and selectivity for different values of γ



Note: z^u is a linearly increasing function of the time-equivalent migration costs π^u for uneducated agents; the figure is drawn for $\gamma_1 < 0 < \gamma_2$, and it represents the evolution of migrants' quality along an iso-migration curve k, as $z^e = g_k(z^u)$, so that a higher value of z^u correspond to a more selective immigration policy.

Figure 1 represents the relationship between migrants' quality and z^u, which is negatively related to the probability of self-selection into migration of uneducated agents, along an iso-migration curve for different values of γ that we have just derived in Proposition 1.[17] When $\gamma = 0$, migrants' quality monotonically increases with z^u along any iso-migration curve. Specifically, quality increases from μ_1^u to μ_1^e when the share of educated agents among the migrants goes from 0 to 1,[18] as the average quality of both educated and uneducated migrants is unaffected by variations in migration costs.

When migrants are negatively selected on unobservables, i.e., $\gamma < 0$, quality is a non-monotonic function of z^u along an iso-migration curve, but the critical point of the function $y[z^u, g_k(z^u)]$, which is implicitly defined by the condition $g_k(z^u) = f(z^u) \equiv z^u - (\mu_1^e - \mu_1^u)/\gamma$, represents a global minimum; migrants' quality monotonically increases with z^u beyond this critical point, and it approaches to its global maximum when the share of educated migrants converges to 1, as shown in Figure 1. Thus, an attempt of admitting only educated migrants is the quality-maximizing choice for the destination country when migrants' wages are a non-increasing function of migration costs, i.e., $\gamma \leq 0$, so that the quality of educated migrants improves with a scale-preserving increase in selectivity.

When migrants are positively selected on unobservables, i.e., $\gamma > 0$, then migrants' quality is maximized when $z^u = z^u(k)$, with $z^u(k)$ being implicitly defined by the condition $g_k(z^u) = f(z^u)$. Thus, attempting to admit only educated agents at destination is never a quality-maximizing choice in this case. The following Corollary further characterizes the immigration policy that maximizes migrants' quality when $\gamma > 0$:

Corollary 1 *When migrants are positively selected on unobservables, migrants' quality is maximized when the average log wage at destination of the set of educated agents who are indifferent between a foreign and a domestic job coincides with the average log wage at destination of the corresponding set of uneducated agents.*

Proof. The condition that denotes the indifference between a domestic and a foreign job is $\epsilon_{i2} = z^l$, for $l \in \{u, e\}$. The assumption of bivariate normality for ϵ_{i0} and ϵ_{i1} implies that $E\left(\ln w_{i1}^l | \epsilon_{i2} = z^l\right) = \mu_1^l + \gamma z^l$, so that $E\left(\ln w_{i1}^e | \epsilon_{i2} = z^e\right) = E\left(\ln w_{i1}^u | \epsilon_{i2} = z^u\right)$ requires that:

$$\mu_1^e + \gamma z^e = \mu_1^u + \gamma z^u.$$

Moving terms around, this condition can be rewritten as $z^e = f(z^u)$. ■

[17]Notice that a higher value of γ is associated with a higher level of migrants' quality for any (z^u, z^e).

[18]The distributional assumptions on ϵ_{i0} and ϵ_{i1}, which have an infinite support, entail that the share of educated migrants can never attain the value of 0 or 1.

A scale-preserving increase in selectivity that pushes z^u beyond $z^u(k)$ reduces migrants' quality, as it would push the log wage at destination of the marginal educated agents below the corresponding value for the marginal uneducated agents. It is also straightforward to demonstrate that the average log wage for educated migrants is higher than the corresponding average log wage for uneducated migrants when $\gamma > 0$ and $z^u = z^u(k)$.[19]

3.2 Scale of migration and optimal share of educated migrants

A movement along the curve $z^e = f(z^u)$, which identifies all the quality-maximizing pairs of z^u and z^e, induces both a variation in the scale of migration and a variation in the share of educated agents among the migrants. The following Proposition establishes how the quality-maximizing share of educated migrants $\beta[z^u, f(z^u)]$ varies with the scale of migration:

Proposition 2 *The share of educated migrants that maximizes migrants' quality is a decreasing function of the scale of migration when migrants are positively selected on unobservables.*

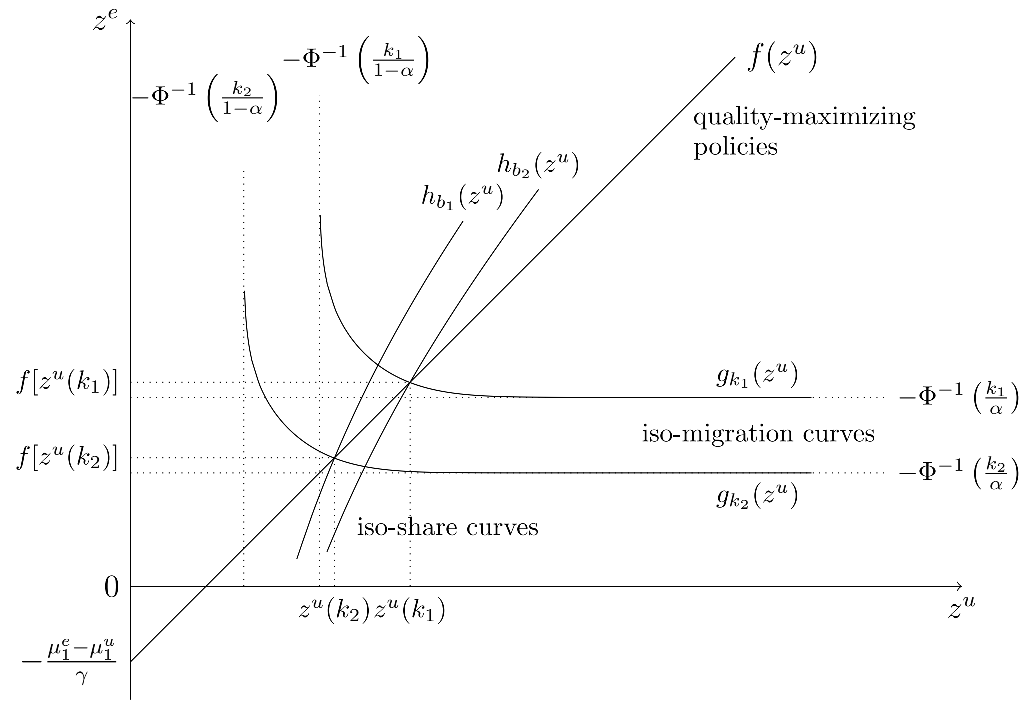
Proof. See Appendix A.2. ■

Proposition 2 demonstrates that a destination country which aims at increasing the scale of the incoming migration flows should let the share of *un*educated migrants increase to ensure that migrants' quality is maximized.[20] Figure 2 provides a graphical representation of this prediction: it plots two different iso-migration curves, with $k_2 > k_1$, and the corresponding pairs of optimal immigration policies, with $z^u(k_2) < z^u(k_1)$ and $f[z^u(k_2)] < f[z^u(k_1)]$, together with the two upward-sloping iso-share curves $z^e = h_{b_1}(z^u)$ and $z^e = h_{b_2}(z^u)$ passing through each of the two pairs of optimal immigration policies. The proof of Proposition 2 hinges on the comparison of the slope of the iso-share curve with the curve $z^e = f(z^u)$ when the two cross: as the curve $z^e = f(z^u)$ is flatter than the iso-share curve $z^e = h_b(z^u)$ in correspondence to their intersection, this implies that an increase in the scale of migration along the curve $z^e = f(z^u)$ leads to a higher iso-share curve, which corresponds to a lower share of educated agents among the migrants.

[19]The proof of this result follows from the fact that the Inverse Mills ratio is a contraction mapping (Heckman, 1979, p. 157).

[20]Notice that migrants' quality necessarily declines with the scale of migration k when migrants are positively selected on unobservables, but this decline is minimized when the share of *un*educated migrants rises.

Figure 2: Quality-maximizing immigration policies for different scales of migration



Note: z^e and z^u are negatively related the difference between the log wage at destination and the log wage at origin, net of migration costs, for educated and uneducated agents respectively; the figure is drawn for $\gamma > 0$, with $k_2 > k_1$ and $b_2 > b_1$, where b_2 and b_1 are two different shares of educated agents among the migrants.

3.3 Relation to the empirical literature

The motivation of our theoretical model resides in a basic empirical fact: observed characteristics, such as education, account for a limited portion of the variance in earnings across individuals. The dispersion of the distribution of earnings at one point in time for a given level of education might be reflecting both the influence of unforecastable time-varying factors and heterogeneity in stable individual-specific characteristics, such as innate ability or talent. Our model implicitly assumes that the latter, which is the dimension of the variance in earnings across potential migrants that destination countries are concerned about, is the factor that generates a dispersion in the earnings of the agents with the same level of education. The empirical relevance of this assumption is corroborated by the survey article by Cunha and Heckman (2007), who conclude that "most variability across people is due to heterogeneity and not uncertainty" (p. 888).

If *un*observed traits play a key role in accounting for the dispersion in earnings for a given set of observed characteristics, then migrants are likely to be a self-selected group also with respect to these traits. The two main theoretical predictions of our model are derived under the assumption that migrants have higher wages at destination than non-migrants with the same level of education, i.e., they are positively selected on unobservables. The empirical evidence on the prevailing pattern of selection on unobservables clearly depends on the "arbitrary nature of the division of earnings into predicted and residual earnings" (Kaestner and Malamud, 2014, p. 89), while it is easier to gather evidence on the pattern of migrants' selection on education, an individual characteristic that can be readily observed in the data. In this respect, a (nearly) universal empirical regularity is that the propensity to migrate is higher among individuals with tertiary education than among less-educated individuals. Indeed, using data on bilateral migrant stocks in OECD destinations in 2000, Docquier *et al.* (2009) show that, on average, 5.5 percent of the individuals with post-secondary education born in a country reside abroad, while the corresponding figure for individuals with less than secondary education stands at 1.3 percent.[21] While the patterns of selection on education and on unobservables can in principle 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). This argument entails that the prevailing pattern of positive migrants' selection on education that is observed in the data could be matched by positive selection on unobservables.[22]

Our model predicts that migrants' quality is maximized when the probability of self-selection into migration for educated agents is higher than the corresponding probability for uneducated agents.[23,24] This, in turn, entails that the effect on migrants' quality of a scale-

[21] Artuç *et al.* (2015) rely on data that cover also non-OECD destinations, and the emigration rate for individuals with tertiary education stands at 8.1 percent, above the rate for less educated individuals.

[22] In line with this argument, 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, while Kaestner and Malamud (2014) uncover different patterns of selection along the two dimensions.

[23] When migrants are positively selected on unobservables, then the maximization of their quality requires that $z^e = f(z^u) < z^u$, which implies that $\Phi(-z^u) < \Phi[-f(z^u)]$.

[24] Notice that a higher emigration rate for educated agents does *not* require the return to education to be larger at destination than at origin, as a greater propensity to migrate among educated agents could be

preserving increase in selectivity when the emigration rate for high-educated individuals is higher than the emigration rate for less-educated individuals is *a priori* ambiguous, while a scale-preserving increase in selectivity would be certainly beneficial in the opposite (but less likely) case of a higher emigration rate for less-educated individuals.

Interestingly, we can also observe that the migration literature has shown that an increase in the size of migration networks at destination is generally associated with a reduction in migration costs that induces both an increase in the scale of migration and a decline in the share of high-educated migrants (see, for instance, McKenzie and Rapoport, 2010 and Beine *et al.*, 2011). Our theoretical model suggests that such a reduction could actually be consistent with the objective of maximizing migrants' quality when the scale of migration expands rather than at odds with it.

4 Extensions

We consider here four extensions of the basic specification of our model that do not alter the theoretical prediction that a scale-invariant increase in selectivity can actually reduce migrants' quality. Specifically, we discuss the implications of (*i*) allowing for a greater dispersion in the quality of educated agents, (*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 Different covariance matrices

The analysis of the model was conducted under the hypothesis that $\mathbf{\Sigma}^e = \mathbf{\Sigma}^u$, which implies that unobservable individual characteristics have the same influence on the log wages of educated and uneducated agents. The productivity of educated workers, which are assigned to more complex tasks, might actually be more sensitive to their ability than the corresponding

induced by the lower time-equivalent migration costs that they face, something that is generally assumed in the literature (Schultz, 1975; Chiquiar and Hanson, 2005; McKenzie and Rapoport, 2010; Beine *et al.*, 2011); Bertoli *et al.* (2013) provide evidence that the time-equivalent migration costs faced by Ecuadorians moving either to the United States or to Spain significantly decline with their level of education.

productivity of uneducated workers, that perform more basic tasks.[25] To address this issue, consider the more general assumption that $\mathbf{\Sigma}^e = a^2\mathbf{\Sigma}^u$, for $a \geq 1$.[26] It is straightforward to show that this assumption implies that:

$$\gamma^e = \frac{(\sigma_1^e)^2 - \rho_{01}\sigma_0^e\sigma_1^e}{[(\sigma_1^e)^2 + (\sigma_0^e)^2 - 2\rho_{01}\sigma_0^e\sigma_1^e]^{1/2}} = a\gamma^u. \tag{9}$$

If both types of migrants are positively selected on unobservables, i.e., $\gamma^e, \gamma^u > 0$, then (9) entails that $\gamma^e \geq \gamma^u$. This, in turn, implies that the intensity of selection on unobservables for a given probability of self-selection into migration–and its responsiveness to a change in this probability–is stronger for educated than for uneducated migrants. This does not alter the prediction of Proposition 1, as we still have that migrants' quality evolves in a non-monotonic way along an iso-migration curve. Pushing selectivity too far still entails that the destination country will induce to self-select into migration a set of educated agents whose log wage falls below the log wage of the set of uneducated agents who are discouraged from migrating. Thus, we have that the maximization of migrants quality requires, as described in Corollary 1, that the average log wage at destination of the set of educated agents, who are indifferent between a foreign and a domestic job, coincides with the average log wage at destination of the corresponding set of uneducated agents. This leads to:

$$z^e = f(z^u) \equiv \frac{\gamma^u}{\gamma^e} z^u - \frac{\mu_1^e - \mu_1^u}{\gamma^e}. \tag{10}$$

Proposition 2, which demonstrates that the quality-maximizing share of educated agents among the migrants is a decreasing function of the scale of migration, is also robust when we introduce assumptions on the two covariance matrices that entail that $\gamma^e = a\gamma^u$, with $a \geq 1$. Its proof hinges on the comparison of the slopes of the curve $z^e = f(z^u)$, which identifies optimal immigration policies, and of the iso-share curve $z^e = h_b(z^u)$ at their intersection. The slope of the iso-share curve is unaffected by the introduction of a more general assumption on the two covariance matrices, as we still have that:

$$\frac{\partial h_b(z^u)}{\partial z^u} = \frac{\lambda(z^u)}{\lambda(z^e)}.$$

[25]For instance, Chen (2008) demonstrates that the variance of residual of log earnings increases with the level of education of the workers, and this higher variance partly reflects unobserved heterogeneity across individuals.

[26]This assumption implies a higher variance of log wages for educated agents, while maintaining that $\rho_{01}^e = \rho_{01}^u = \rho_{01}$; the sign of the difference between γ^e and γ^u is, in general, undetermined if $\rho_{01}^u \neq \rho_{01}^e$.

As (10) implies that $f(z^u) < z^u$ and as the Inverse Mills ratio is a monotonically increasing function, we have that the slope of iso-share curve $h_b(z^u)$ is higher than 1 when this crosses the curve $z^e = f(z^u)$. It is straightforward to see from (10) that the slope of this latter curve is equal to $\gamma^u/\gamma^e = 1/a \leq 1$, and this concludes the proof. Thus, both our main predictions are robust when allowing for a greater variance in the log wages of educated agents: attempting to admit only educated agents at destination is detrimental for migrants' quality, and the quality-maximizing share of educated migrants is negatively related to the scale of migration.

4.2 Random variation in time-equivalent migration costs

The basic specification of the model retains 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 by including 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:

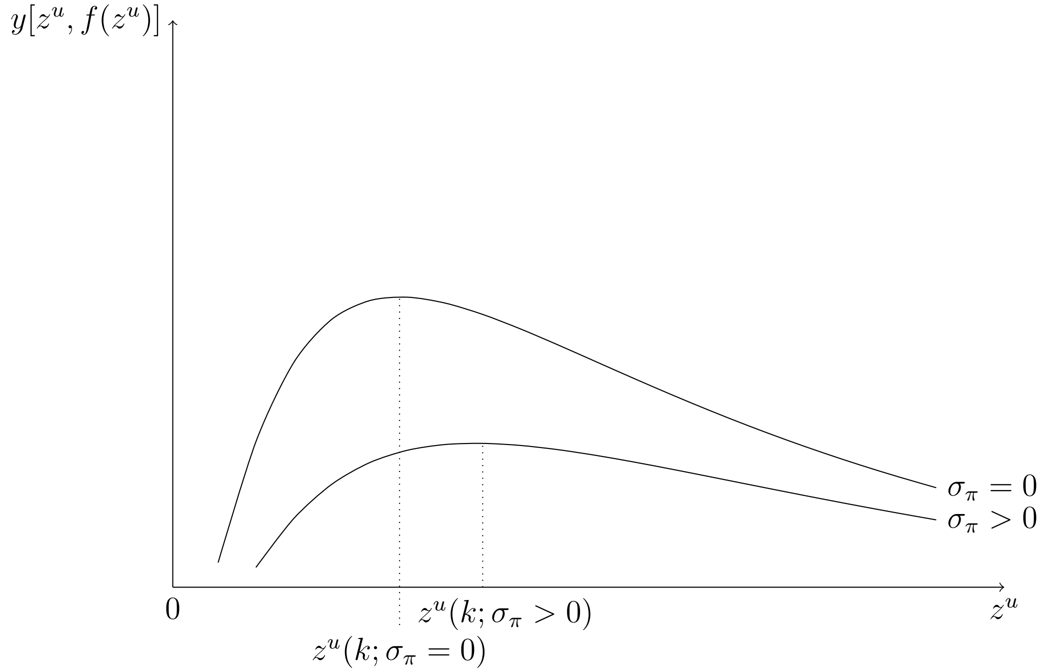
$$\Pr\left(\widetilde{\epsilon}_{i2} \equiv \epsilon_{i1} - \epsilon_{i0} - \epsilon_{i\pi} > \mu_0^l + \mu_\pi^l - \mu_1^l\right) = \Phi(-\widetilde{z}^l),$$

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

$$\widetilde{Q}_1(\widetilde{z}^l) = \left(\widetilde{\gamma} - \frac{\sigma_{1\pi}}{\widetilde{\sigma}_2}\right)\lambda(\widetilde{z}^l). \tag{11}$$

In the absence of covariance between the stochastic component of the log wage at destination and the stochastic component of migration costs, i.e., $\sigma_{1\pi} = 0$, then we obtain the same pattern of selection on unobservables than in Section 2, but with $|\widetilde{Q}_1(\widetilde{z}^l)| < |Q_1(z^l)|$ as $|\widetilde{\gamma}| < |\gamma|$ and $\widetilde{z}^l < z^l$. Intuitively, self-selection on noneconomic factors 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., $\widetilde{Q}_1(\widetilde{z^l}) > 0$, as the indirect adverse effect of the policy change becomes weaker, as depicted in Figure 3. Self-selection on non-economic factors also does not affect the result derived in Proposition 2, as it only changes the responsiveness of migrants' quality with respect to a scale-preserving variation in migration costs.

Figure 3: Unobserved heterogeneity in preferences for migration



Note: the figure represents the evolution of migrants' quality along an iso-migration curve 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 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 (11), which determines the pattern of migrants' self-selection on unobservables, can be positive even if $\widetilde{\gamma} \leq 0$, when the correlation between the 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 our prediction that migrants' quality can decline with a scale-preserving increase in selectivity, introducing heterogeneity in preferences for migration would strengthen our theoretical prediction under the empirically relevant assumption that $\rho_{1\pi} < 0$.

4.3 An alternative informational structure

We have assumed that wages are *remotely* observable as in Borjas (1987, 1999b), 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:[27,28]

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

where

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

Bertoli (2010b) demonstrates that:

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

where:

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

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 a positive selection on unobservables compared

[27] Migrants have domestic wages belonging to the lower (upper) tail of the truncation of $\ln w_{i0}^l$ when $\rho_{01} > \sigma_0/\sigma_1$ ($\rho_{01} < \sigma_0/\sigma_1$).

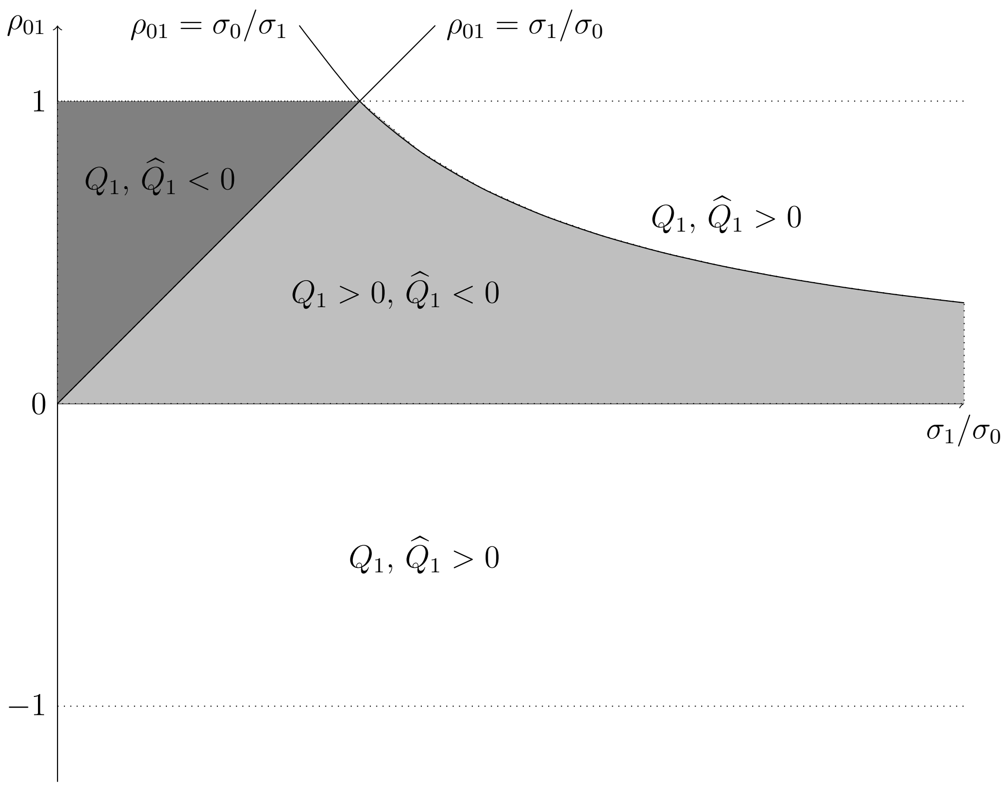
[28] The probability of self-selection into migration is always lower when wages are only locally rather than remotely observable if less than half of the population at origin migrates, i.e., $\mu_0 + \pi - \mu_1 > 0$. This follows from the fact that $[(\sigma_1\rho_{01}/\sigma_0) - 1]\sigma_0 < \sigma_2$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

to Borjas (1987), as depicted in Figure 4, but it does not affect our theoretical predictions: a scale-invariant increase in selectivity can reduce migrants' quality when migrants' are positively selected on unobservables, and the quality-maximizing share of educated migrants is inversely related to the scale of migration. Specifically, when wages are only locally observable and positively selected on unobservables, then the maximization of migrants' quality requires that:

$$z^e = \widehat{f}(z^u) \equiv z^u - \frac{\mu_1^e - \mu_1^u}{\widehat{\gamma}}.$$

As discussed in the next section, the sign of the difference between γ and $\widehat{\gamma}$ depends, in general, on the elements of the covariance matrix $\mathbf{\Sigma}$ and on the scale of migration, so that it is, in general, not possible to sign the differential between $f(z^u)$ and $\widehat{f}(z^u)$.

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



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 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 the one in Bertoli (2010b) to that of Borjas (1987), 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.[29]

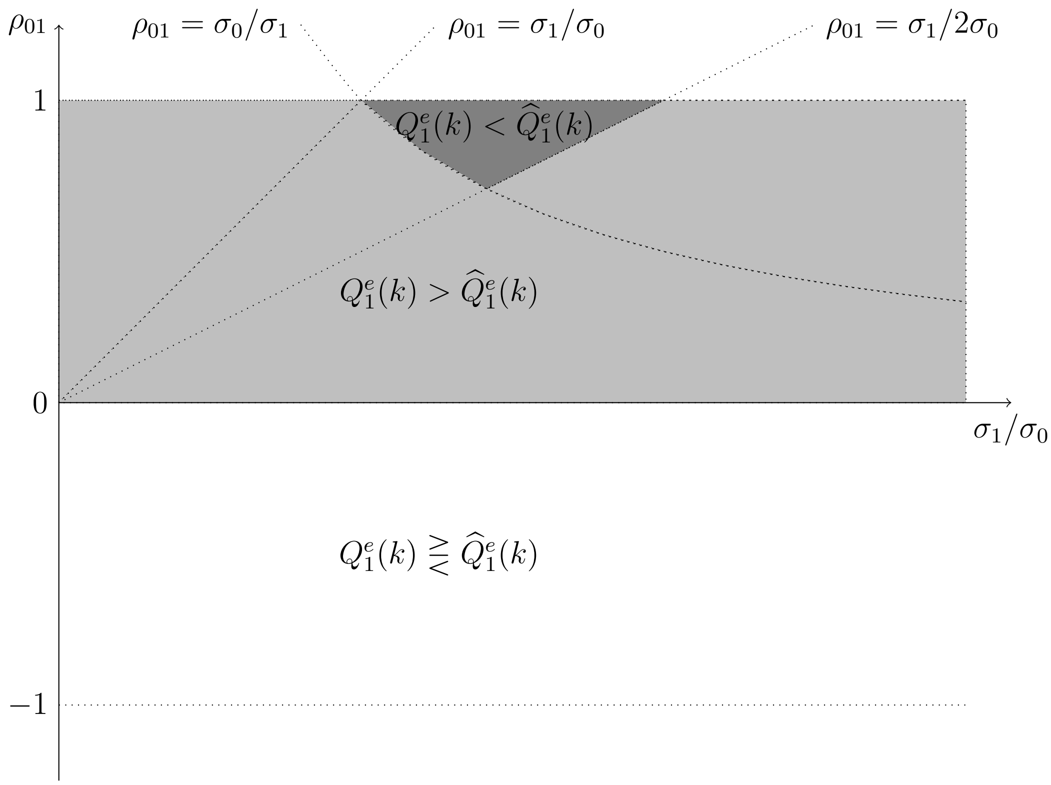
We have that better information *reduces* migrants' quality when $\rho_{01} > \sigma_0/\sigma_1$ and $\rho_{01} > \sigma_1/2\sigma_0$ or when $\rho_{01} < 0$ if the scale k of migration is sufficiently small,[30] as depicted in Figure 5. Remarkably, the proposed change in the informational structure is detrimental for migrants' quality when unobservable skills can be easily transferred across countries, i.e., ρ_{01} 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.

Once this change in the informational structure (for educated agents) takes place, does this alter our theoretical predictions concerning the non-monotonicity of migrants' quality along an iso-migration curve and the relationship between the quality-maximizing share of educated migrant and the scale of migration? We can focus on the case where (i) the destination country has an interest in implementing this change in the informational structure, i.e., migrants' quality unambiguously increases, and (ii) both types of migrants are positively selected on unobservables, as assumed in Corollary 1 and Proposition 2. Both conditions are

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

[30] See Appendix A.3 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 $\mu_0^e + \pi^e > \mu_1^e$.

met when ρ_{01} is such that $\sigma_1/2\sigma_0 > \rho_{01} > \sigma_0/\sigma_1$, as can be seen from Figures 4 and 5. In this case, we have that:[31]

$$\gamma^e \equiv \frac{\sigma_1^2 - \sigma_{01}}{\left(\sigma_1^2 + \sigma_0^2 - 2\sigma_{01}\right)^{1/2}} > \frac{\sigma_{01}}{\sigma_0} \equiv \gamma^u.$$

We know (from Section 4.1 above) that this represents a sufficient condition for demonstrating that our main theoretical predictions hold.

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

[31]See Appendix A.3 for the demonstration of this inequality.

policy change influences migrants' selection on unobservables, such as ability and motivation, which contribute to determine 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. Increasingly selective immigration policies might not just be "unfriendly to development" (Pritchett, 2006), but they might also fail to attain their main goal of raising migrants' quality pursued by recipient countries. Furthermore, an expansion in the share of *un*educated agents among the migrants could be in the self-interest of a destination country that is expanding the scale of incoming migration flows. 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.

Artuç, E., F. Docquier, Ç. Özden and C. Parsons (2015): "A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations," *World Development*, 65, 6–26.

Aydemir, A. (2011): "Immigrant Selection and Short-Term Labor Market Outcomes by Visa Xategory," *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.

Belot, 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., 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*, 117(2), 565–591.

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.

——— (2015): "The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again," *Journal of Human Capital*, forthcoming.

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.

Chen, S. (2008): "Estimating the Variance of Wages in the Presence of Selection and Unobserved Heterogeneity," *Review of Economics and Statistics*, 90(2), 275–289.

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.

Cunha, F., and J. J. Heckman (2007): "Identifying and Estimating the Distributions of *Ex Post* and *Ex Ante* Returns to Schooling," *Labour Economics*, 14, 870–893.

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*, 117(2), 303–346.

Fernández-Huertas Moraga, J. (2011): "New Evidence on Emmigrant 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.

Heckman, J. J. and T. Kautz (2012): "Hard Evidence on Soft Skills," *Labour Economics*, 19(4), 451–464.

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., and H. Rapoport (2010): "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *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.

Pritchett, L. (2006): *Let Their People Come*. Washington: Center for Global Development.

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

The partial derivative of migrants' quality in (4) with respect to z^u along an iso-migration curve is given by:

$$\begin{aligned}\frac{\partial y[z^u, g_k(z^u)]}{\partial z^u} &= \frac{\partial \beta[z^u, g_k(z^u)]}{\partial z^u}\left[\mu_1^e + Q_1[g_k(z^u)] - \mu_1^u - Q_1(z^u)\right] \\ &+\beta[z^u, g_k(z^u)]\left[\frac{\partial Q_1[g_k(z^u)]}{\partial z^u} - \frac{\partial Q_1(z^u)}{\partial z^u}\right] + \frac{\partial Q_1(z^u)}{\partial z^u}.\end{aligned}$$

We can rewrite using (8) this partial derivative as follows:

$$\begin{aligned}\frac{\partial y[z^u, g_k(z^u)]}{\partial z^u} &= -\frac{\partial g_k(z^u)}{\partial z^u}\lambda[g_k(z^u)]\beta[z^u, g_k(z^u)]\left[\mu_1^e + Q_1[g_k(z^u)] - \mu_1^u - Q_1(z^u)\right] \\ &+\beta[z^u, g_k(z^u)]\left[\frac{\partial Q_1[g_k(z^u)]}{\partial z^u} - \frac{\partial Q_1(z^u)}{\partial z^u}\right] + \frac{\partial Q_1(z^u)}{\partial z^u}.\end{aligned}$$

As:

$$\frac{\partial Q_1(z^l)}{\partial z^l} = \gamma\lambda(z^l)[\lambda(z^l) - z^l],$$

we can rewrite it once more as follows:

$$\begin{aligned}\frac{\partial y[z^u, g_k(z^u)]}{\partial z^u} &= -\frac{\partial g_k(z^u)}{\partial z^u}\lambda[g_k(z^u)]\beta[z^u, g_k(z^u)]\Big[\mu_1^e - \mu_1^u + \gamma\lambda[g_k(z^u)] - \gamma\lambda(z^u)\Big] \\ &+\gamma\beta[z^u, g_k(z^u)]\left[\frac{\partial g_k(z^u)}{\partial z^u}\lambda[g_k(z^u)](\lambda[g_k(z^u)] - g_k(z^u)) - \lambda(z^u)(\lambda(z^u) - z^u)\right] \\ &+\gamma\lambda(z^u)[\lambda(z^u) - z^u].\end{aligned}$$

If $\gamma = 0$, then this expression simplifies to:

$$\frac{\partial y[z^u, g_k(z^u)]}{\partial z^u} = -\frac{\partial g_k(z^u)}{\partial z^u}\lambda[g_k(z^u)]\beta[z^u, g_k(z^u)]\Big(\mu_1^e - \mu_1^u\Big) > 0,$$

which entails that migrants' quality monotonically increases with z^u along an iso-migration curve. The monotonicity follows from the fact that migrants' wages are independent from migration costs. When $\gamma \neq 0$, then migrants' quality increases with z^u along an iso-migration curve if and only if:

$$-\gamma\left[\frac{\partial g_k(z^u)}{\partial z^u}\lambda[g_k(z^u)]\left[\frac{\mu_1^e - \mu_1^u}{\gamma} + g_k(z^u) - \lambda(z^u)\right] - \frac{1 - \beta[z^u, g_k(z^u)]}{\beta[z^u, g_k(z^u)]}\lambda(z^u)[\lambda(z^u) - z^u]\right] > 0.$$

As:

$$\frac{\partial g_k(z^u)}{\partial z^u} = \frac{\beta[z^u, g_k(z^u)] - 1}{\beta[z^u, g_k(z^u)]} \frac{\lambda(z^u)}{\lambda[g_k(z^u)]} < 0,$$

with some tedious but straightforward algebra the inequality above simplifies to:

$$\gamma \left[g_k(z^u) - f(z^u) \right] > 0,$$

where:

$$f(z^u) \equiv z^u - \frac{\mu_1^e - \mu_1^u}{\gamma}.$$

When γ is higher (lower) than zero, then $y[z^u, g_k(z^u)]$ monotonically increases (decreases) with z^u along an iso-migration curve when $g_k(z^u) > f(z^u)$, while it monotonically decreases (increases) with z^u when $g_k(z^u) < f(z^u)$.

A.2 Proof of Proposition 2

Let $z^e = h_b(z^u)$ be a family of iso-share curves, indexed by b, which gives the unique value of z^e such that $\beta[z^u, h_b(z^u)] = b$. From (5), we have that:

$$h_b(z^u) = -\Phi^{-1} \left[\frac{(1-\alpha)b}{\alpha(1-b)} \Phi(-z^u) \right].$$

Deriving $h_b(z^u)$ with respect to z^u, and exploiting the rule of the derivation of an inverse function, we get:

$$\frac{\partial h_b(z^u)}{\partial z^u} = \frac{(1-\alpha)b}{\alpha(1-b)} \frac{\phi(-z^u)}{\phi(-z^e)}.$$

Substituting b with $\beta(z^u, z^e)$ from (5):

$$\frac{\partial h_b(z^u)}{\partial z^u} = \frac{(1-\alpha)\frac{\alpha\Phi(-z^e)}{\kappa(z^u,z^e)}}{\alpha\left(1 - \frac{\alpha\Phi(-z^e)}{\kappa(z^u,z^e)}\right)} \frac{\phi(-z^u)}{\phi(-z^e)}.$$

With simple algebraic manipulations, and recalling that the definition of the scale of migration $\kappa(z^u, z^e)$ in (6), we get:

$$\frac{\partial h_b(z^u)}{\partial z^u} = \frac{(1-\alpha)\frac{\Phi(-z^e)}{\kappa(z^u,z^e)}}{1 - \frac{\alpha\Phi(-z^e)}{\kappa(z^u,z^e)}} \frac{\phi(-z^u)}{\phi(-z^e)} == \frac{\Phi(-z^e)}{\Phi(-z^u)} \frac{\phi(-z^u)}{\phi(-z^e)};$$

As $\lambda(z^l) \equiv \phi(-z^l)/\Phi(-z^l)$, for $l \in \{u, e\}$, we eventually get:

$$\frac{\partial h_b(z^u)}{\partial z^u} = \frac{\lambda(z^u)}{\lambda(z^e)}.$$

When migrants are positively selected on unobservables, we know that migrants' quality is maximized for $z^e = f(z^u) < z^u$; as the Inverse Mills ratio is a monotonically increasing function, this implies that $\lambda(z^u) > \lambda[f(z^u)]$, and this in turn entails that:

$$\frac{\partial h_b(z^u)}{\partial z^u} > 1$$

when a iso-share curve crosses the curve that identifies the quality-maximizing combinations of z^u and z^e. We also know that $\partial f(z^u)/\partial z^u = 1$; hence, a joint reduction in z^u and z^e along this curve, which determines an increase in the scale of migration, results in a reduction in the quality-maximizing share of educated agents among the migrants.

A.3 Change in the informational structure and migrants' quality

A scale-preserving change in the informational structure, with remotely observable wages for educated individuals, increases educated migrants' quality, i.e., $Q_1^e(k) > \widehat{Q}_1^e(k)$, if and only if:

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

where $\widehat{z}^e$ gives rise to a scale of migration equal to k under the informational structure in Bertoli (2010b). We have that (A.1) clearly holds when $\rho_{01} < \min\{\sigma_1/\sigma_0, \sigma_0/\sigma_1\}$, as this entails that $\widehat{\gamma} < 0 < \gamma$. When $\rho_{01} > \sigma_0/\sigma_1$, then (A.1) can be rewritten as:

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

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

$$\sigma_0^2\sigma_1^2(\sigma_1^2 - 2\sigma_{01}) > \sigma_{01}^2\left(\sigma_1^2 - 2\sigma_{01}\right). \tag{A.2}$$

If $\rho_{01} < \sigma_1/2\sigma_0$, then (A.2) 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 (A.2) 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:

$$\frac{\sigma_{12}}{\sigma_2} \gtreqless -\frac{1-\Phi(\widehat{z}^e)}{\Phi(\widehat{z}^e)}\frac{\sigma_{01}}{\sigma_0}, \tag{A.3}$$

unless we introduce assumptions on the value of $\widehat{z}^e$. Specifically, (A.3) 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:

$$\frac{\sigma_{12}}{\sigma_2} > -\frac{\sigma_{01}}{\sigma_0} > -\frac{1-\Phi(\widehat{z}^e)}{\Phi(\widehat{z}^e)}\frac{\sigma_{01}}{\sigma_0}, \tag{A.4}$$

when $\widehat{z}^e > 0$. Again, we have that the first inequality in (A.4) 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:

$$\sigma_0^2\sigma_1^2(\sigma_1^2 - 2\sigma_{01}) < \sigma_{01}^2\left(\sigma_1^2 - 2\sigma_{01}\right). \tag{A.5}$$

As $\rho_{01} > \sigma_1/2\sigma_0$, then (A.5) 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.