Optimal Design of an Immigration Points System

John McHale and Keith Rogers

Queen’s School of Business
Queen’s University
Kingston, ON, K7L 3N6
Canada

jmchale@business.queensu.ca
krogers@business.queensu.ca

Draft: November 2005

We gratefully acknowledge financial support from Social Science and Humanities research Council of Canada (File # 537-2004-1006) under the INE Skills Research Initiative. We also thank Mohsen Bouaissa of Human Resources and Skills Development Canada his valuable comments and advice on the paper. All errors and omissions are our own.
Abstract
This paper explores the design of an immigration points system based on two simple elements: a human-capital based earnings regression for making error-minimizing predictions of immigrant success in the host labor market, and a predicted-earnings threshold for selecting whom to accept and reject. We derive what we call the selection frontier, which captures how the average earnings (or “quality”) of the admitted immigrant pool declines with the number of accepted immigrants as lower thresholds are adopted. We show how the position of the frontier is determined by the accuracy of earnings predictions as well as the mean and variance of post-entry earnings in the applicant pool. We also examine how the frontier is affected by the policies of “competitor” countries in the market for skilled immigrants. An illustrative example is developed to explore how the predicted-earnings threshold is optimally chosen given policy makers willingness to trade off immigrant quality for quantity.

Note: This paper is the first part of a two-part sequence that develops and applies design principles for a skills-based points system. This part develops the basic design principles and evaluative tools and is primarily theoretical. The second empirical part will apply the methods developed here to the design and evaluation of the Canada’s points system.
1. Introduction

There is growing interest in industrialized countries in skill-focused immigration policies. Many policy makers believe that higher-earning immigrants add more value in the labor market, are net contributors to the fiscal system, assimilate better into host-country societies, and pose a lesser threat of crime and terrorism. A number of longer-term trends appear to be driving the shift to more skill-focused approaches, notably population aging, skill-biased technical change, and the complementary effects of broader globalization.

There are many examples of this shift towards more skill-focused policies, even if the steps are still sometimes tentative. Canada has revamped its points system to make it more focused on human capital based indicators of long-term success in the labor market, and has been increasing the share of the skilled stream in its overall immigrant flow. Australia has introduced a major new class of long-stay temporary visas aimed at the highly skilled and has also been increasing the share of its skills stream for permanent migration. The U.K. has introduced and extended a points-based Highly Skilled Migration Programme on a pilot basis, and both the Labour and Conservative parties offered proposals for fully fledged points systems as part of their 2005 election manifestos. After extensive debate, Germany adopted a more skill focused system at the beginning of 2005, but backtracked on earlier plans to adopt a fully fledged points system. The U.S. greatly expanded the availability of temporary H-1B visas for the

---

1 For one thing, population aging will increasingly place retirement income and health care systems under fiscal strain. Although immigration does not provide a “silver bullet” for solving these problems, governments will have a strong incentive to admit high earnings individuals who place little demand on state transfer programs and services, especially where these individuals do not stay long enough to acquire entitlement to retirement-related benefits.

2 In recent decades, skill-biased technical change has been increasing the gap in earnings and employment opportunities between less-educated and more educated workers. Besides any effects on the relative economic attractiveness of less and more skilled workers, these labor market effects have important implications for the political economy of immigration policy. When employers of the highly skilled are facing rising wage costs, they have more incentive to lobby for relaxations on restrictions on the recruitment of foreign skilled workers. And when highly skilled workers are experiencing strong absolute and relative wage growth, they have less incentive to resist such recruitment.

3 In theory, international trade and capital flows can substitute international migration; in practice, the three forms of integration tend to occur together. The classic modern case is the EU. Increased competitive pressure in global industries is one factor pushing employers to seek access to foreign talent and whatever competitive advantage they might bring. The broader forces of globalization are also creating a global elite, with members that find it easy to settle in any of the world’s cosmopolitan centers.
highly skilled in the late 1990s and early 2000s. Although the bursting of the high tech bubble (and post-September 11, 2001 security concerns) undermined the constituency for renewing the expanded cap after it expired in 2003, an under-supply of skilled workers is already leading to calls to reduce restrictions on the recruitment of foreign talent.

With this level of policy interest, it is surprising that the question of the optimal design of a skill-focused points system has not received more attention. This paper explores the optimal design question based on a very simple idea: a points system can be devised based on a human capital-based earnings regression for predicting how potential immigrants will “perform” in the domestic labor market and a chosen threshold for predicted earnings for deciding who to accept and who to reject. We show that these two elements are sufficient to determine the optimal point allocations for various bundles of human capital characteristics. The resulting flexible framework is also amenable to ex ante and ex post normative analysis of the immigrant selection system. Of course, the idea of giving points for observed human capital characteristics is the essential feature of existing points systems. While the allocations of points in the Canadian and Australian systems are clearly informed by the findings from the human capital literature, the allocation process appears to lack a firm analytical foundation. We hope this paper will provide that foundation.

Having high predicted earnings is certainly a rather narrow basis for selecting immigrants. Most countries also place importance on reunifying families and protecting people fleeing persecution or humanitarian catastrophes. Even from a narrow economic perspective of those already present in the host country, a better measure of immigrant value is the “surplus” that the country gains from the immigrants. This surplus can be defined as the value the country receives less what they must pay to the immigrants. Simple models show that it is not necessarily the most highly skilled immigrants that generate the greatest surplus. However, the likelihood that the surplus increases with the skills of the migrant tends to rise when we allow for fiscal effects, knowledge spillovers,

---

4 The cap was expanded from 65,000 to 195,000 visas per year.
5 See, for example, National Science Board (2003).
6 See, for example, Borjas (1995), and McHale (2003).
or the value of specialized skills. Augmenting the relative supply of skilled workers should also reduce overall earnings inequality, so that skilled recruitment can be desirable on both efficiency and equity grounds. But whatever the merits of focusing narrowly on skills, it is the case that a number of countries are striving to select more skilled immigrant pools. It is thus worthwhile to look for a more systematic approach to designing a skills-based selection system.

We determine the “optimal” points system in two stages. The first stage is the identification of the selection frontier. The frontier is optimal in the sense that the average earnings of the admitted pool is maximized for any given pool size, and is found by using the earnings regression that minimizes prediction errors. Each point on the frontier maps to a unique predicted-earnings threshold. The second stage is the identification of the optimal threshold. This threshold obviously depends on policy-maker preferences over immigrant quality and quantity, which in turn depend on the effects of immigration on the economy and broader society. Rather than offer some unavoidably special model of the effects of immigration, we simply posit a willingness to tradeoff quality for quantity based on the idea that policy makers like immigrant human capital but face a convex cost of adjusting to immigration. We see the main contribution of the paper as being the first-stage identification of the frontier and present the second stage simply to illustrate the nature of the choice facing the policy maker.

The rest of the paper is structured as follows. In Section 2, we begin with a one-period horizon and certain immigrant employment to show the basic mechanism for designing a linear points system for a given predicted-earnings threshold. We then show how the basic set-up can be extended to allow for more realistic multi-year horizons,

---

7 George Borjas (1999, p. 19) makes the following case for focusing on immigrant skills: If nothing else, decades of social science research have established an irrefutable link between human capital—a person’s endowment of ability and acquired skills—and a wide array of social and economic outcomes, ranging from earnings potential to criminal activity, work effort to drug abuse, and from family stability to life expectancy. In view of this strong link, it is not surprising that the United States cares about whether the immigrant population is composed of skilled or unskilled workers. The skill composition of the immigrant population—and how the skills compare to those of natives—determine the social and economic consequences of immigration for the country. [Emphasis in the original.]
immigrant assimilation, and an earnings threshold measured in terms of the present discounted value of the immigrant’s predicted earnings stream. Finally, we discuss the possibility of non-employment and the implications of relaxing the requirement that the points system is linear.

In Section 3, we derive the selection frontier and show how its location depends on the mean and variance of the (lognormally distributed) applicant pool and also the variance of the prediction error for the earnings regression. We model the admitted pool as a “selected sample” in the well defined sense that the pool is an incidentally truncated sub-sample of the applicant pool. The truncation takes place on the basis of a comparison of predicted earnings and the chosen predicted earnings threshold. In Section 4, we look to how the skill-targeting policies of “competitor” countries in the international market for mobile talent affects the distribution of earnings in a country’s applicant pool and thereby the position of its selection frontier. In Section 5, we use a simple example of policy-maker preferences to show how the predicted-earnings threshold is determined as the most-preferred point on the selection frontier. Section 6 has some concluding comments and outlines our future research directions.

2. Basic Points System Design

In this section, we describe how a points system can be designed using a human capital-based earnings regression for predicting applicant “success” (i.e. predicted earnings, $\hat{Y}$) in the host-country labor market combined with a designated predicted earnings threshold, $\hat{Y}^*$. As noted in the introduction, the key idea is that applicants with predicted earnings above the threshold are accepted; all others rejected. We initially restrict attention to linear points systems, which in turn restricts us to using additively separable functional forms for the earnings regression. (A linear system gives the total points by simply adding up the points per unit of each human capital characteristic.) The earnings regression is additively separable if there is some monotonically increasing transformation that yields a right-hand side that is linear in the variables.
(i) One-period horizon and certain employment

To focus on the essential elements, we start with a very simple earnings regression, certain employment, and a one-period horizon for time spent in the host-country labor market. There are just two human capital indicators for a potential immigrant, \( i \): years of schooling \((S_i)\) and years of experience \((E_i)\), and the additive separable earnings regression is assumed to take the familiar log-linear (or semi-log) form,

\[
\ln Y_i = y_i = \beta_0 + \beta_1 S_i + \beta_2 E_i + u_i \quad u_i \sim n(0, \sigma_u^2).
\]

The equation is estimated by OLS, yielding an equation for what we assume is the best-linear-unbiased predictor of earnings for potential immigrant \( i \),

\[
\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 S_i + \hat{\beta}_2 E_i.
\]

Equation (2) is an equation for predicted log earnings \((\hat{y}_i)\). However, what we need is the log of predicted earnings \((\ln \hat{Y}_i)\), which we obtain using the approximation,

\[
\ln \hat{Y}_i \approx \frac{\hat{\sigma}_u^2}{2} + \hat{y}_i = \frac{\hat{\sigma}_u^2}{2} + \hat{\beta}_0 + \hat{\beta}_1 S_i + \hat{\beta}_2 E_i.
\]

We now slightly rearrange the equation (and henceforth ignore the approximation) to obtain,

\[
\ln \hat{Y}_i = \frac{\hat{\sigma}_u^2}{2} - \hat{\beta}_0 = \hat{\beta}_1 S_i + \hat{\beta}_2 E_i.
\]

Finally, we force the left-hand side of \((3')\) to equal an arbitrarily chosen 100 points when predicted earnings exactly equals the predicted-earnings threshold, \(\hat{Y}_i = \hat{Y}^*\).
\[ \frac{100\bar{\beta}}{\bar{\beta}} = \left( \frac{100\hat{\beta}_1}{\bar{\beta}} \right) S_i + \left( \frac{100\hat{\beta}_2}{\bar{\beta}} \right) E_i, \]

(4)

where \( \bar{\beta} = \ln \hat{Y}^* - \frac{\bar{\sigma}_e^2}{2} - \hat{\beta}_0. \)

The coefficients on the schooling and experience variables give the number of points that should be granted per unit of schooling and experience respectively. Applicants who score 100 points or more (or equivalently have predicted earnings greater than the threshold, \( \hat{Y}^* \)) are accepted; applicants who score less than 100 points are rejected. The combinations of schooling and experience that result in exactly 100 points are shown by the boundary line in Figure 1. The relative value of an additional year of schooling in terms of years of experience is given by the slope of the boundary line. The figure also shows which bundles of human capital characteristics lead to acceptance and which lead to rejection.

**Figure 1**

Experience, \( E_i \)

Acceptance Set (includes boundary line for individuals with exactly 100 points)

Rejection Set

Slope = \( \frac{-\hat{\beta}_1}{\hat{\beta}_2} \)

Schooling, \( S_i \)
Clearly, the success of such a points system depends on the predictive success of the earnings regression. As is well known, prediction errors can result biased estimators of the coefficients, sampling error in the estimated coefficients, measurement error in the explanatory variables, and random disturbances in actual earnings. Figure 2 graphically shows how prediction errors will lead to “mistakes” in the immigrant screening process. In Section 3, we make the simplifying assumptions that the true earnings regression is known and the human capital variables that the immigration authorities observe are measured without error. Thus the first three sources of error mentioned above are conveniently not present, and we focus on the variance of the disturbance term (which under these simplifying assumptions is equal to the variance of the prediction error) as the key determinant of the success of the screening process for any given applicant pool and predicted earnings threshold.

**(ii) Allowing for a multi-period horizon**

An obvious limitation of this simple model is that immigrants will typically be present in the host-country labor market for longer than a single year. This will force the policy maker to consider a threshold for the present discounted value of the predicted earnings stream rather than a single year’s predicted earnings. Predicted earnings for
later years will be affected by how the immigrants earnings evolve with time spent in the host-country labor market. The simplest possible modification of our basic case is to add a years-since migration \((t_i)\) to our basic framework, so that the earnings regression becomes,

\[
(5) \quad \ln Y_{it} = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 t_i + u_{it} \quad u_{it} \sim n(0, \sigma_u^2).
\]

Again using OLS to obtain the best-linear-unbiased predictor of earnings and the approximation used above, we can write the log of predicted earnings as,

\[
(6) \quad \ln \hat{Y}_{it} \approx \hat{\gamma}_{it} + \frac{\hat{\sigma}^2_u}{2} = \hat{\beta}_0 + \hat{\beta}_1 S_i + \hat{\beta}_2 E_i + \hat{\beta}_3 t_i + \frac{\hat{\sigma}^2_u}{2}.
\]

Assuming a time horizon of \(T_i\) years and a discount rate of \(\delta\), we use (6) to write the present discounted value of earnings as,

\[
(7) \quad \hat{Z}_i = \int_0^{T_i} e^{-\delta t} \hat{Y}_{it} dt_i
= e^{\frac{\hat{\sigma}^2_u}{2} + \hat{\gamma}_{it} + \hat{\beta}_1 S_i + \hat{\beta}_2 E_i} \int_0^{T_i} e^{(\hat{\beta}_3 - \delta) t_i} dt_i
= e^{\frac{\hat{\sigma}^2_u}{2} + \hat{\gamma}_{it} + \hat{\beta}_1 S_i + \hat{\beta}_2 E_i} \left( \frac{1}{\hat{\beta}_3 - \delta} \right) (e^{(\hat{\beta}_3 - \delta) T_i} - 1).
\]

We can preserve linearity in the points system if approximate the last term in parentheses by \(e^{(\hat{\beta}_3 - \delta)}\). (That is, we simply ignore the fact that 1 is subtracted from this term, which should be a reasonable approximation for longer horizons). Using this approximation and taking logs yields,

\[
\ln \hat{Z}_i \approx \frac{\hat{\sigma}^2_u}{2} + \hat{\gamma}_{it} + \hat{\beta}_1 S_i + \hat{\beta}_2 E_i - \ln(\hat{\beta}_3 - \delta) + (\hat{\beta}_3 - \delta) T_i.
\]
We further assume that the immigrant will work until age $A$, so that $T_i = A - A_i$, where $A_i$ is age at arrival. Now letting our threshold for the present discounted value of the predicted earnings stream equal $\hat{Z}^*$ (and again imposing a points cut off of 100 and ignoring the approximations), we can write our key points equation as,

$$\frac{100\hat{\beta}}{\hat{\beta}} = \left( \frac{100\hat{\beta}_1}{\hat{\beta}} \right) S_i + \left( \frac{100\hat{\beta}_2}{\hat{\beta}} \right) E_i - \left( \frac{100\hat{\beta}_3}{\hat{\beta}} \right) A_i,$$

(8)

where $\hat{\beta} = \ln \hat{Z}^* - \frac{\hat{\sigma}_u^2}{2} - \hat{\beta}_0 + \ln (\hat{\beta}_3 - \hat{\delta}) - (\hat{\beta}_3 - \hat{\delta})A_i$.

Once again, the terms in parentheses on the right-hand-side provide the points per unit of the human capital characteristic. The key difference between (8) and (4) is that (8) allows for a points penalty based on the age of the applicant (reflecting the fact that older applicants will have fewer years of earnings in the host-country labor market).

(iii) Allowing for non-employment

Up to this point, we have assumed all immigrants are employed in the host economy and have focused on developing a log-linear model for predicted earnings conditional on that employment. However, immigrant non-employment is likely to be a significant concern for various reasons—lengthy job search in an unfamiliar labor market, non-recognition of immigrant credentials, poorly transferable skills, etc. If we continue to use predicted earnings as the basis for selection, the obvious extension to our model is to treat predicted (unconditional) earnings $(\hat{Y}_\mu^u)$ as the product of the predicted probability of employment $(\hat{J}_\mu^u)$ and predicted earnings conditional on employment $(\hat{Y}_\mu^e)$. Taking logs, we have an amended predicted earnings equation: $\ln \hat{Y}_\mu^u = \ln \hat{J}_\mu^u + \ln \hat{Y}_\mu^e$.

Based on the vast literature on empirical earnings functions, we have argued that a log-linear specification is defensible for modeling the determinants of $\ln \hat{Y}_\mu$. This is what allowed us develop a user-friendly linear points system. But additive separability is much harder to defend for the probability of employment. If, for example, we
let \( J_{it} = e^{\alpha_0 + \alpha_1 S_i + \alpha_2 E_i} \), then taking logs of both sides clearly gives us the necessary linear right-hand side:
\[
\ln J_{it} = \alpha_0 + \alpha_1 S_i + \alpha_2 E_i.
\]
But a log-linear specification is unlikely to be defensible for modeling the probability of employment. One obvious limitation is that the probability of employment can exceed unity. Of course, more standard models of this probability, such as probit or logit, do not produce the necessary linearity after the log transformation. This forces us to balance the costs of restrictive functional forms against the simplicity of linearity.\(^8\) We next turn to develop a framework that allows us to explore how improvements the underlying fit of the earnings-prediction model affect the performance of the immigrant selection system.

### 3. Derivation of the Selection Frontier

In this section we derive what we call the selection frontier facing policy makers based on the design principles described above. This selection frontier captures the tradeoff between the average earnings (or "quality") of the selected pool and the number (or "quantity") of immigrants admitted from a given applicant pool. The position of this frontier provides an easy to interpret method of evaluating the performance of the selection system.

In deriving the frontier, a key simplification is that we assume the actual (post-entry) earnings are lognormally distributed across the applicant pool. In addition, we assume that the earnings prediction errors that result from the earnings regression are lognormally distributed. Together these assumptions imply that actual earnings and

---

\(^8\) Note also that even though the use of the log-linear functional form is standard for earnings functions (conditional on employment), our requirement of additive separability rules out interaction and higher order terms (e.g. experience squared) as additional explanatory variables. Given the ubiquity of such terms in estimated earnings regressions, such exclusions are likely to be quite restrictive. Thus we may wish to abandon the simplicity of a linear points system even without the complication of the possibility of non-employment. This raises the broader question of how important it is that the points system is linear. Clearly, a linear system is user friendly, as potential applicants can easily understand how their points total is arrived at, and also what they would need to do to increase their score. On the other hand, even highly non-linear systems can be made reasonably user friendly through the use of an on-line points calculator. The potential applicant could enter a set of characteristics, and the calculator would reveal the points score based on the predicted earnings for someone with those characteristics. The decision about adopting a non-linear system would involve a tradeoff between lost simplicity and improved predictions. The systematic approach to points-system design we develop in this paper has the virtue of allowing designers to explore how predictions are improved when non-linearities are allowed.
predicted earnings have a joint lognormal distribution with a correlation coefficient equal to the square root of the coefficient of determination \((R^2)\) from the earnings regression.\(^9\)

The first step in deriving the frontier is to determine the expected earnings of the admitted pool for any given earnings threshold. To do this, we treat the admitted pool as resulting from the incidental truncation of the applicant pool, where the applicant is admitted if their predicted earnings, \(\hat{Y}_i\), is greater than the predicted earnings threshold, \(\hat{Y}^*\) (where \(i\) subscripts are now dropped for notional convenience). The expected earnings of the admitted pool then given by,

\[
E[Y | \hat{Y} \geq \hat{Y}^*] = \frac{\int E[Y | \hat{y} = x] \, f(x) \, dx}{\int P(x \leq \hat{y}^*) \, dx}.
\]

We apply this general specification to our model by substituting in the relationship between the conditional estimate of earnings given the estimator of log earnings that was introduced in the previous section: \(E[Y | \hat{y} = x] \approx e^{\sigma^2/2} e^x\). We also substitute the normal probability density function (PDF) for the general PDF to yield the conditional expectation:

\[E[Y | \hat{Y} \geq \hat{Y}^*] = \frac{\int e^{\sigma^2/2} e^x \, f(x) \, dx}{\int 1 - P(x \leq \hat{y}^*) \, dx}.
\]

\(^9\) The normality of the distribution of predicted log earnings can be demonstrated by using moment generating functions (MGF). Using notation described in the main text the relationship between log earnings, predicted log earnings and the residual is:

\[y_i = \hat{y}_i + u_i\]

By construction, the estimated log income is independent of the residual. As a result, the MGF for the sum of these variables equals the product of the left hand side variables’ respective MGFs:

\[MGF(y) = MGF(\hat{y})MGF(u)\]

Since \(y_i \sim n(\overline{y}, \sigma_{y}^2)\) and \(u_i \sim n(0, \sigma_{u}^2)\) we can substitute in the specific moment generating functions:

\[
\exp\left(\overline{y}t + \frac{\sigma_{y}^2 t^2}{2}\right) = MGF(\hat{y})\exp\left(\frac{\sigma_{u}^2 t^2}{2}\right)
\]

Re-arranged we obtain:

\[MGF(\hat{y}) = \exp\left(\overline{y}t + \frac{(\sigma_{y}^2 - \sigma_{u}^2)t^2}{2}\right),\]

which demonstrates that \(\hat{y} \sim n(\overline{y}, \sigma_{y}^2 - \sigma_{u}^2)\).
(10) \( E[Y \mid \hat{y} \geq \hat{y}^*] = \frac{1}{P(\hat{Y} \geq \hat{Y}^*)} \int_{\hat{y}^*}^{\infty} \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{(x - \mu_y)^2}{2\sigma_y^2}} e^x dx \)

Collecting like terms and completing the square yields:  

\[
(11) \quad = \frac{e^{\sigma_y^2/2}}{1 - \Phi(Z_y(\hat{y}^*))} \int_{\hat{y}^*}^{\infty} \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\left[\frac{x-(\sigma_y^2+\mu_y)}{\sigma_y}\right]^{2}/2} e^{-\frac{-(\sigma_y^2+\mu_y)^2}{2\sigma_y^2}} dx 
\]

\[
= \frac{e^{\sigma_y^2/2}}{1 - \Phi(z(\hat{y}^*))} \int_{\hat{y}^*}^{\infty} e^{\sigma_y^2/2} e^{-\left[\frac{x-(\sigma_y^2+\mu_y)}{\sigma_y}\right]^{2}/2} dx 
\]

By defining \( \xi = (\hat{y}^* - \sigma_y^2) \) and using a change of variables to \( \nu = (x - \sigma_y^2) \) the second part of this expression becomes:

\[
(12) \quad \int_{\hat{y}^*}^{\infty} e^{-\left[\frac{x-(\sigma_y^2+\mu_y)}{\sigma_y}\right]^{2}/2} dx = \int_{\xi}^{\infty} e^{\sigma_y^2/2} e^{-\left[\frac{(\gamma)}{\sigma_y}\right]^{2}/2} d\nu 
\]

Substituting this result into (11) yields:

\[
(13) \quad E[Y \mid \hat{y} \geq \hat{y}^*] = \frac{e^{\sigma_y^2/2}}{1 - \Phi(z(\hat{y}^*))} \int_{\hat{y}^*}^{\infty} e^{\sigma_y^2/2} e^{-\left[\frac{(\gamma)}{\sigma_y}\right]^{2}/2} d\nu 
\]

Integrating and substituting in the standard normal transformation, \( z_y(k) = \left(\frac{k - \mu_y}{\sigma_y}\right) \) yields:

---

\[ \text{See Appendix B for omitted steps.} \]
Collecting terms this simplifies to:

\[
E[Y \mid \hat{y} \geq \hat{y}^*] = e^{-\frac{1}{2} \sigma_j^2} \left( \frac{1 - \Phi(z_j (\hat{y}^* - \sigma_j^2))}{1 - \Phi(z_j (\hat{y}^*))} \right)
\]

Since \( e^{\frac{1}{2} \sigma_j^2} = e^{\frac{1}{2} \sigma_j^2} e^{\mu_j} \) we have:

\[
E[Y \mid \hat{y} \geq \hat{y}^*] = e^{\frac{1}{2} \mu_j} \left( \frac{1 - \Phi(z_j (\hat{y}^* - \sigma_j^2))}{1 - \Phi(z_j (\hat{y}^*))} \right)
\]

For the final step we use the fact that \( z_j (\hat{y}^* - \sigma_j^2) = z_j (\hat{y}^* - \sigma_j^2) \) and equation (14) to obtain:

\[
(15) \quad E[Y \mid \hat{y} \geq \hat{y}^*] = \bar{y} \left( \frac{1 - \Phi(z_j (\hat{y}^* - \sigma_j^2))}{1 - \Phi(z_j (\hat{y}^*))} \right)
\]

With this parsimonious equation for the expected earnings of the admitted pool, we can determine how the “quality” of the admitted pool changes with the predicted-earnings threshold.
Proposition 1: The expected earnings of the admitted pool is increasing in the predicted – earnings threshold.¹¹

Proof:

Many of our proofs will follow the form of calculating a derivative of the conditional expectation function and then characterizing its sign. In this case, we will characterize the derivative of the expected earnings with respect to the cutoff earnings:

\[
\frac{\partial E[Y \mid \hat{y} \geq \hat{y}^*]}{\partial \hat{y}^*} = \frac{\phi(z_{\hat{y}}(\hat{y}^*) - \sigma_{\hat{y}}))(1 - \Phi(z_{\hat{y}}(\hat{y}^*) + \phi(z_{\hat{y}}(\hat{y}^*))(1 - \Phi(z_{\hat{y}}(\hat{y}^*) - \sigma_{\hat{y}}))}{\sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*)))^2}
\]

Since the average earnings of the candidate pool, \(\bar{Y}\) and the denominator of the fraction, \(\sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*))^2\) are both positive, the sign depends on the numerator of the fractional component. Specifically, the expected earnings of the admitted pool will increase in the earnings threshold iff:

\[
\phi(z_{\hat{y}}(\hat{y}^*) - \sigma_{\hat{y}}))(1 - \Phi(z_{\hat{y}}(\hat{y}^*)) < \phi(z_{\hat{y}}(\hat{y}^*))(1 - \Phi(z_{\hat{y}}(\hat{y}^*) - \sigma_{\hat{y}}))
\]

Since \(\phi(\cdot)\) is the standard normal PDF and \(\Phi(\cdot)\) is the standard normal cumulative distribution function (CDF), this condition can be expressed in terms of normal hazard rate functions. The average earnings of successful candidates will increase if:

\[
\frac{\phi(z_{\hat{y}}(\hat{y}^*) - \sigma_{\hat{y}})}{1 - \Phi(z_{\hat{y}}(\hat{y}^*) - \sigma_{\hat{y}})) < \frac{\phi(z_{\hat{y}}(\hat{y}^*))}{(1 - \Phi(z_{\hat{y}}(\hat{y}^*)))^2}
\]

¹¹ See Appendix A for additional details on the proof.
Since the normal hazard rate is an increasing function, we know this condition must hold and therefore expected earnings of the admitted pool increases in the predicted earnings threshold. □

The second step in deriving the selection frontier is to determine the relationship between the share of immigrants admitted and the predicted earnings threshold. We define the share of candidates admitted as
\[ p = \frac{N}{T}, \]
where \( N \) is the number admitted and \( T \) is the total pool size. Candidates are admitted when their predicted earnings exceeds the predicted-earnings threshold, \( \hat{Y}^* \). Given the lognormal distribution of earnings in the applicant pool, the share of applicants admitted for any given threshold can be written as:

\[
(18) \quad p = 1 - \Phi(z_{\hat{y}}(\hat{Y}^*)).
\]

**Proposition 2:** The share of the applicant pool that is admitted is decreasing in the predicted-earnings threshold.

**Proof:**

Again, we will find and characterize the appropriate derivative. In this case, the change in the proportion admitted given a change in predicted earnings threshold. Taking the derivative from (18) we obtain:

\[
(19) \quad \frac{\partial p}{\partial \hat{Y}^*} = -\phi(z_{\hat{y}}(\hat{Y}^*)) \frac{\hat{Y}^*}{\sigma_y}.
\]

Since \( \phi(z_{\hat{y}}(\hat{Y}^*)) \) is the normal PDF, it is positive as are the cutoff income, \( \hat{Y}^* \) and the standard deviation of the distribution of incomes \( \sigma_y \). Since all the variables are positive, the negative sign in front forces the overall function to be negative confirming that a rise in earnings threshold reduces the proportion of the applicant pool that is admitted. □
We are now in a position to derive the selection frontier. Since \( z_y(\cdot) \) and \( \Phi(\cdot) \) are positive monotonic functions, we can invert them and make minor rearrangement of (18) to obtain:

\[
(20) \quad \hat{y}^* = z_y^{-1}(\Phi^{-1}(1 - p)).
\]

This establishes a one-to-one relationship between a cutoff threshold and rate of share of the applicant pool that is admitted. We next subtract \( \sigma_y^2 \) from both sides and transform both sides by \( z_y \) to obtain:

\[
(21) \quad z_y(\hat{y}^* - \sigma_y^2) = \Phi^{-1}(1 - p) - \sigma_y^2.
\]

Recalling that \( z_y(\hat{y}^*) - \sigma_y = z_y(\hat{y}^* - \sigma_y^2) \), we now substitute (18) and (21) into (15) to obtain an equation for the selection frontier.

\[
E[Y \mid p] = \frac{\bar{Y}}{p}(1 - \Phi(z_y^{-1}(\Phi^{-1}(1 - p)) - \sigma_y^2)))
\]

Working through the formulas for \( z_y^{-1} \) and \( z_y \) we obtain:

\[
E[Y \mid p] = \frac{\bar{Y}}{p}(1 - \Phi(\frac{\sigma_y\Phi^{-1}(1 - p) + \mu_y - \sigma_y^2 - \mu_y}{\sigma_y}))
\]

Which simplifies to:

\[
(22) \quad E[Y \mid p] = \frac{\bar{Y}}{p}(1 - \Phi(\Phi^{-1}(1 - p) - \sigma_y))).
\]

The selection frontier is shown in Figure 3 for the case where \( \bar{Y} = 65,000, \sigma_y^2 = 0.4, \) and \( \sigma_y^2 = 0.32 \). Note that the latter two parameters imply that the \( R^2 \) from the
earnings regression is equal to 0.2. The frontier shows how the “quality” of the admitted pool declines as a larger share of the applicant pool is admitted, where it is assumed that the best possible means of predicting earnings is being utilized. As expected, as we move towards admitting all applicants, the expected earnings of the admitted pool converges to the mean earnings of the applicant pool.

We next establish that the selection frontier in downward sloping, so that the admission of larger share of the applicant pool (which is achieved by lowering the earnings threshold) is associated with a decline in the expected earnings of the admitted pool.

*Proposition 3: The expected earnings of the admitted pool is decreasing in the share of applicant pool admitted.* \(^{12}\)

Proof:

\(^{12}\) See Appendix A for additional details on the proof.
This time we take the derivative of the selection frontier function with respect to price and show the sign is everywhere negative.

\[
\frac{\partial E[Y \mid p]}{\partial p} = -\frac{\bar{Y}}{p} \left( \frac{1 - \Phi(\Phi^{-1}(1 - p) - \sigma_j)}{p} + \frac{\phi(\Phi^{-1}(1 - p) - \sigma_j)}{\phi(\Phi^{-1}(1 - p))} \right)
\]

Average earnings, \(\bar{Y}\) and proportion of candidates admitted, \(p\) are positive, so the sign of the equation is determined by sum of the fractions in brackets. The first part of this term is positive because its numerator is 1 minus an outcome of a CDF which is bounded above by 1 and its denominator is the share of candidates admitted which is positive. The second component is the ratio of two normal PDFs which are everywhere positive. As a result, the sum of these terms is positive so that the negative sign at the front renders the expression is negative. Thus we have the result that the conditional expectation of earnings decreases as \(p\) increases. □

The selection frontier determines the quality-quantity tradeoff available to policy makers. We next examine the factors that determine the position of frontier. Given our lognormality assumptions, the position of the frontier is determined by just three parameters: the mean earnings of the applicant pool (\(\bar{Y}\)), the variance of earnings in the applicant pool (\(\sigma^2_y\)), and the variance of the prediction error (\(\sigma_u^2\)).

**Proposition 4:** The selection frontier is shifted upwards by an increase in the pool’s mean earnings.

\[
\frac{\partial E[Y \mid p]}{\partial \bar{Y}} = \frac{1}{p} (1 - \Phi(\Phi^{-1}(1 - p) - \sigma_j))
\]

Proof:
Because \( p \) is positive and the \( \Phi(\cdot) \) function is bounded above by 1, the value of this derivative is positive for all values of \( p \). Thus an increase in \( \bar{Y} \) results in higher earnings for accepted candidates at each value of \( p \) and is therefore associated with an upward shift of the selection frontier. \( \square \)

**Proposition 5:** The selection frontier is shifted upwards by an increase in the variance of earnings in the applicant pool and shifted downwards by an increase in the variance of the prediction error.\(^{13} \)

**Proof:**

Again we find the derivative with respect to the standard deviation term and characterize its sign. In this case, we start with the derivative for the expected income of successful candidates condition on a rate of admissions \( p \).

\[
\frac{\partial E[Y | p]}{\partial \sigma_y} = \frac{\bar{Y}}{p} \left( \phi(\Phi^{-1}(1-p) - (\sigma_y^2 - \sigma_u^2)^{1/2}) \right) \frac{\sigma_y}{(\sigma_y^2 - \sigma_u^2)^{1/2}}
\]

\[
+ \frac{\partial \bar{Y}}{\partial \sigma_y} \left( 1 - \Phi(\Phi^{-1}(1-p) - (\sigma_y^2 - \sigma_u^2)^{1/2}) \right)
\]

This equation is the sum of two terms; we will prove this sum is positive by showing that each of the components of the sum is positive.

The component is the multiplication of several elements: earnings \( \bar{Y} \) and proportion of candidates accepted \( p \) are clearly positive while \( \phi(\cdot) \) is the standard normal PDF and is also positive. The final part must be positive if the earnings estimation explains any of the variance, \( \sigma_y^2 - \sigma_u^2 = \sigma^2 \) will be positive and since \( \sigma_y \) is a standard deviation, it too is positive. In short, the first component is positive.

\(^{13}\) See Appendix A for additional details on the proof.
The second component includes the relationship between the partial change in pool earnings and pool variance. This is given by \( \frac{\partial \bar{Y}}{\partial \sigma_y} = \frac{\sigma_y \bar{Y}}{2} \) which is greater than 0. The rest of the term is 1 minus a point under the normal CDF which is bounded above by 1 and therefore the whole term is greater than 0.

Since both components of the sum are positive, the total must be positive and as a result, the average earnings for successful candidates increase with an increase in variance.

Repeating this analysis for the derivative with respect to variance of the prediction error \( \sigma_u \) is quite a bit simpler:

\[
(26) \quad \frac{\partial E[Y | p]}{\partial \sigma_u} = -\frac{\bar{Y}}{p} \left( \phi(\Phi^{-1}(1-p) - \sigma_y^2) \right) \frac{\sigma_u}{\sigma_y^2}
\]

Since income \( \bar{Y} \) proportion admitted \( p \) are positive \( \phi(\cdot) \) is a normal PDF and the ratio of two standard errors must be positive, the leading negative sign renders the entire expression negative. Consequently, the average earnings of successful candidates falls with an increase in prediction error. □

Figure 4 shows the selection frontier for different values of the \( R^2 \) from the earnings regression (recalling that \( R^2 = 1 - \frac{\sigma_u^2}{\sigma_y^2} \)). A reduction in the variance of the prediction error will raise the \( R^2 \) and shift the selection frontier upwards. From the graph we can see that the increase in expected earnings due to better prediction is greater the lower the proportion of the applicant pool admitted. In other words, the benefits of better prediction increase as the system becomes more selective.
4. International Competition for Talent and the Position of the Selection Frontier (Incomplete)

In this section, we consider how developments in competitor countries in the market for immigrants could affect the position of the selection frontier facing a country’s policy makers. In terms of the model developed above, the key question is how these developments affect the distribution of earnings in the applicant pool. One promising approach would be to apply the classic Roy model to examine the decisions of would-be immigrants to self select into the applicant pool given the alternatives available to them in other countries (including their home country). In this section, however, we explore a simpler example with two immigrant destinations, which for concreteness we take to be Canada and the United States. As a further simplification, we make the extreme assumption that the United States is the preferred choice for all immigrants. More realistically, we take it the United States uses an employer-driven approach to recruiting skilled immigrants and further assume that the probability of being accepted in
the United States is an increasing function of actual earnings. Actual earnings are taken to be equal in both labor markets. All immigrants not admitted to the United States are assumed to become part of the applicant pool for Canada. Our focus is on how policy developments in the United States affect the position of Canada’s selection frontier.

Our modeling approach builds on techniques developed for a different purpose in Deaton (2005). Deaton’s interest is in the income distribution of individuals participating in a survey when the probability of participation declines in a known way with income. Our interest is in distribution of earnings of the applicant pool to Canada when the probability of being in that pool—i.e. the probability of not being accepted as an immigrant to the United States—declines in a known way with earnings. Deaton assumes that the initial distribution is lognormal, and develops the assumptions required for the distribution of actual respondents to be approximately lognormal. This technique is useful in the context of our model, since it allows Canada’s applicant pool to be lognormally distributed even after the United States has selected certain individuals from its originally lognormally distributed applicant pool, and thus allows us to apply the framework developed in Section 3. The United States selection system is captured by the relationship between the probability of selection and potential immigrant’s earnings. A more successful selection system will do a better job of “creaming off” high earnings individuals, which we will see has implications for the position of Canada’s selection frontier.

More specifically, we assume that the probability of an applicant being rejected by the United States is 1 for log earnings below a threshold value, $y^\ast$, and rises at an increasing rate with the log earnings above that value. This “creaming off” by the United States changes the density of the applicant pool to Canada from $f(y)$ to $\tilde{f}(y)$, where the latter is given by,

$$\tilde{f}(y) = \frac{p^*(y)f(y)}{\int p(s)dF(s)} = \frac{p^*(y)}{\bar{p}} f(y),$$

(27)
where \( \bar{p} \) and \( p(y) \) are the average and marginal probabilities that someone will be available to Canada:

\[
(28) \quad p(y) = \begin{cases} 
1 & \text{for } y \leq \mu_y - \theta \sigma_y \\
-\alpha (y - \mu_y + \theta \sigma_y) - \frac{\gamma}{2} (y - \mu_y + \theta \sigma_y)^2 & \text{for } y \geq \mu_y - \theta \sigma_y 
\end{cases}
\]

Where \( \theta = \left( \frac{y' - \mu_y}{\sigma_y} \right) = -z_y(y') \) and \( \alpha \) and \( \gamma \) are positive parameters that capture how the probability of being selected by the United States varies with earnings.

When this probability is applied to the original density function \( f(y) \) we obtain:

\[
(29) \quad \tilde{f}(y) = \begin{cases} 
\frac{1}{\bar{p} \sigma_y \sqrt{2\pi}} e^{-\frac{(y - \mu_y)^2}{2 \sigma_y^2}} & \text{for } y \leq \mu_y - \theta \sigma_y \\
\frac{1}{\bar{p} \sigma_y \sqrt{2\pi}} e^{-\frac{(y - \mu_y)^2}{2 \sigma_y^2}} e^{-\alpha (y - \mu_y + \theta \sigma_y) - \frac{\gamma}{2} (y - \mu_y + \theta \sigma_y)^2} & \text{for } y \geq \mu_y - \theta \sigma_y 
\end{cases}
\]

As shown in Appendix C, provided \( \theta \) is large enough, the post-competition density function \( \tilde{f}(y) \) for log earnings is associated with an approximately normal distribution with mean \( \tilde{\mu}_y \) and variance \( \tilde{\sigma}_y^2 \):

\[
(30) \quad \tilde{\mu}_y = \mu_y - \frac{\sigma_y^2 (-\alpha + \sigma_y \theta \gamma)}{(1 + \gamma \sigma_y^2)}.
\]

\[
(31) \quad \tilde{\sigma}_y^2 = \frac{\sigma_y^2}{(1 + \gamma \sigma_y^2)}.
\]
We can now examine how an increase in the effectiveness of United States in “creaming off” high earnings individuals (as measured by an increase in $\alpha$ and/or $\gamma$) affects the position of Canada’s selection frontier. It is easy to see that increase in $\alpha$ will lower $\tilde{\mu}$, but leave $\tilde{\sigma}$ unchanged. The result is a decrease in the (arithmetic) mean of earnings in Canada’s applicant pool, and a consequent downward shift of the selection frontier. The affect of an increase in $\gamma$ is more complicated. From equation (31), we see that an increase in $\gamma$ will increase the variance of log earnings, which by itself would lead to a downward shift in the selection frontier. An increase in $\gamma$ will also lead to a reduction in $\tilde{\mu}$, provided that $\theta$ is sufficiently large, which results in a further downward shift of the frontier.

5. Choice of the Predicted-Earnings Threshold: An Illustrative Example

Our focus to this point has been the determination of the selection frontier facing policy makers. We now turn briefly to the question of how policy makers should choose the point on the frontier to operate on; that is, how they choose the predicted-earnings threshold. The chosen point will depend on the willingness of policy makers to trade off immigrant “quality” for “quantity,” which in turn will depend on the details of how skilled immigration affects the economy and politics of recruiting foreign workers. Rather than attempt to provide a detailed model of the determinants of policy-maker preferences, we limit ourselves here to an illustrative example of the trade offs that are likely to be involved. The essence of the example is that policy makers like immigrant human capital but face a convex adjustment in adjusting to immigration.

Letting $q$ denote the average quality (as measured by expected earnings) of the immigrant pool and $N$ the total number of immigrants admitted, we can write the total human capital of the admitted pool as $q \times N$. We assume that policy makers place a value of $a$ on a unit of human capital. There is also a cost to immigration that is a convex function of the number of immigrants: $(b/2)N^2$. Total policy maker utility is thus given by,
\( U = aqN - \frac{b}{2}N^2. \)

The resulting policy-maker indifference curves in quality-quantity space will be U-shaped (see Figure 5). This means that over a certain range a fall in quality can be compensated by an increase in quantity; however, once the level of immigration reaches a certain level, further immigration must be compensated by an increase in quality. More specifically, the slope of an indifference curve through a given point is given by,

\[
\frac{dq}{dN} = \frac{bN - aq}{aN}.
\]

An indifference curve will reach its minimum point when the numerator is equal to zero. Thus the set of minimum points as we move to higher and higher indifference curves will rise along the linear schedule,

\[
q = \frac{a}{b} N.
\]

Figure 5 shows how policy makers choose the predicted-earnings threshold by reaching the highest feasible indifference given the selection frontier.\(^{14}\) [Comparative statics to follow.]

\(^{14}\) In appendix D we show that there exists a unique interior solution to the first order conditions of this problem.
6. Concluding Comments

This paper has explored a simple idea for designing a more rational skills-based immigration points system. This idea is to systematically use the information available from state-of-art human capital earnings regressions to predict a potential immigrant’s success and to choose predicted-earnings threshold below which applicants are rejected. In contrast to our proposed approach, the findings from human capital research have heretofore been applied to selection system design in a seemingly ad hoc way. Not surprisingly, the accuracy of earnings predictions turns out to be a key determinant of the performance of the selection system.

By examining how changes in the specification of the underlying earnings prediction model affect the performance of the selection system, we hope that our framework can provide a means of evaluating various ideas for improving skills-based selection. These ideas include: allowing differential points allocations for domestic and
foreign acquired education and experience;\textsuperscript{15} providing significant point bonuses for applicants with job offers or who are already in employment under a temporary visa program;\textsuperscript{16} considering objective information on the quality of human capital provision in the applicant’s country-of-origin in addition to applicant-specific information;\textsuperscript{17} and weighing the labor market prospects of spouses (and possibly even children) in admission decisions.\textsuperscript{18}

We plan in the next stage of our research to apply our framework to the design and evaluation of a points system for Canada. The central challenge will be the estimation of an immigrant earnings regression that provides good predictions of immigrant labour market success. We hope that this application will allow us to test the workability of our approach to optimal design and to assess of proposed reforms to Canada’s points system.

\textsuperscript{15} Friedberg (2002), for example, documents large differences in returns to domestic and foreign human capital among immigrants to Israel.
\textsuperscript{16} Grubel (2005) argues strongly for emphasizing job offers in the selection of immigrants to Canada.
\textsuperscript{17} Mattoo, Neagu, and Ozden (2005) document sizeable country-of-origin effects for given observed education levels for immigrants to the U.S.
\textsuperscript{18} Card (2005), for example, documents the strong labor market performance of the children of immigrants despite the less than complete earnings convergence of their parents.
Appendix A – Details on Proofs

Proposition 1: The expected earnings of the admitted pool is increasing in the predicted – earnings threshold.

Starting with the formula for average earnings conditional on admission:

\[ E[Y | \hat{y} \geq \hat{y}^*] = \frac{1 - \Phi(z_{\hat{y}}(\alpha) - \sigma_{\hat{y}}))}{1 - \Phi(z_{\hat{y}}(\hat{y}^*)))} \]

\[ \frac{\partial E[Y | \hat{y} \geq \hat{y}^*]}{\partial \hat{y}^*} = \frac{-\phi(z_{\hat{y}}(\hat{y}^*)))}{\sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*)))} + \frac{\phi(z_{\hat{y}}(\hat{y}^*)))}{\sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*)))^2} \] (A2)

Since all other terms are positive, the sign of this depends on the numerator. The average earnings of the pool will decrease with an increase in the threshold if the following condition is met:

\[ \phi(z_{\hat{y}}(\hat{y}^*))) - \sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*))) > \phi(z_{\hat{y}}(\hat{y}^*))) - \sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*))) \] (A4)

This is equivalent to:

\[ \frac{\phi(z_{\hat{y}}(\hat{y}^*))) - \sigma_{\hat{y}}(1 - \Phi(z_{\hat{y}}(\hat{y}^*)))}{1 - \Phi(z_{\hat{y}}(\hat{y}^*)))} > \frac{\phi(z_{\hat{y}}(\hat{y}^*)))}{(1 - \Phi(z_{\hat{y}}(\hat{y}^*)))} \] (A5)

Which is a condition on the hazard rate function of the normal distribution. Since the hazard rate function of the normal distribution is strictly increasing, this condition is never met. Therefore the earnings of the admitted pool must increase in the admissions threshold.

Proposition 3: The expected earnings of the admitted pool is decreasing in the share of applicant pool admitted

\[ E[Y | p] = \frac{Y}{p} (1 - \Phi(\Phi^{-1}(1 - p) - (\sigma_{\hat{y}}^2 - \sigma_u^2)^{1/2})) \]
(A7) \[ \frac{\partial E[Y | p]}{\partial p} = -\frac{\bar{Y}}{p^2} (1 - \Phi(\Phi^{-1}(1 - p) - \sigma_\delta)) - \frac{\bar{Y}}{p} (\phi(\Phi^{-1}(1 - p) - \sigma_\delta) D_p(\Phi^{-1}(1 - p))) \]

By the inverse function theorem:

(A8) \[ D_p(\Phi^{-1}(1 - p)) = \frac{1}{\Phi'(\Phi^{-1}(1 - p))} = -\frac{1}{\phi(\Phi^{-1}(1 - p))} \]

Substituting this in and solving yields:

(A9) \[ \frac{\partial E[Y | p]}{\partial p} = -\frac{\bar{Y}}{p} \left(1 - \Phi(\Phi^{-1}(1 - p) - \sigma_\delta)\right) + \frac{\phi(\Phi^{-1}(1 - p) - \sigma_\delta)}{\phi(\Phi^{-1}(1 - p))} \]

Proposition 5: The selection frontier is shifted upwards by an increase in the variance of earnings in the applicant pool and shifted downwards by an increase in the variance of the prediction error.

(A10) \[ E[Y | p] = \frac{\bar{Y}}{p} (1 - \Phi(\Phi^{-1}(1 - p) - \sigma_\delta)) \]

Substituting in:

(A11) \[ \sigma_\delta = (\sigma^2_y - \sigma^2_u)^{1/2} \]

(A12) \[ E[Y | p] = \frac{\bar{Y}}{p} (1 - \Phi(\Phi^{-1}(1 - p) - (\sigma^2_y - \sigma^2_u)^{1/2})) \]

(A13) \[ \frac{\partial E[Y | p]}{\partial \sigma_u} = -\frac{\bar{Y}}{p} (\phi(\Phi^{-1}(1 - p) - (\sigma^2_y - \sigma^2_u)^{1/2})) \frac{\sigma_u}{(\sigma^2_y - \sigma^2_u)^{1/2}} \]

(A14) \[ = -\frac{\bar{Y}}{p} (\phi(\Phi^{-1}(1 - p) - \sigma_\delta^2) \frac{\sigma_u}{\sigma_\delta} \]

This is negative for all values of \( p \), which means the policy curve becomes increasingly restrictive as prediction error increases.

The partial with respect to \( \sigma_y \) has the added complication that changes in \( \sigma_y \) result in a change of the mean of the distribution \( \bar{Y} \).

(A15) \[ \frac{\partial E[Y | p]}{\partial \sigma_y} = \frac{\bar{Y}}{p} (\phi(\Phi^{-1}(1 - p) - (\sigma^2_y - \sigma^2_u)^{1/2})) \frac{\sigma_y}{(\sigma^2_y - \sigma^2_u)^{1/2}} \]
\[ + \frac{\partial \bar{Y}}{\partial \sigma_y} (1 - \Phi(\Phi^{-1}(1 - p) - (\sigma_y^2 - \sigma_u^2)^{1/2})) \]

Where \( \frac{\partial \bar{Y}}{\partial \sigma_y} = \frac{1}{2} \sigma_y \bar{Y} \), which is greater than 0.

Therefore, \( \frac{\partial E[Y | p]}{\partial \sigma_y} \) is the sum of positive terms and therefore greater than 0.

Consequently, the selection frontier becomes less restrictive as \( \sigma_y \) increase.
Appendix B – Skipped Steps Between Equations (10) and (11)

Grouping the integrand terms involving $x$ in equation (10) we obtain:

\[ (B1) = e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}} \]

\[ (B2) = e^{-\frac{-(x^2-2\mu_jx+\mu_j^2)}{2\sigma_j^2}} \]

\[ (B3) = e^{-\frac{x^2-2\mu_jx+\mu_j^2-2x\sigma_j^2}{2\sigma_j^2}} \]

\[ (B4) = e^{-\frac{x^2-2x(\mu_j+\sigma_j^2)+\mu_j^2-(\sigma_j^2+\mu_j)^2}{2\sigma_j^2}} \]

\[ (B5) = e^{-\frac{-(x-\mu_j)^2+\mu_j^2-(\sigma_j^2+\mu_j)^2}{2\sigma_j^2}} \]

\[ (B6) = e^{\frac{-(x-\mu_j)^2}{2\sigma_j^2}} e^{-\frac{-\mu_j^2+(\mu_j+\sigma_j)^2}{2\sigma_j^2}} \]

Which can be substituted into (10) to yield (11)
Appendix C – Extension of Deaton’s Result to the Quadratic Case\(^\text{19}\)

As described in the main text, the density function for the log earnings of candidates available to Canada is:

\[
\tilde{f}(y) = \begin{cases} 
\frac{1}{\bar{p}\sigma_y \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(y-y_0)^2}{\sigma_y^2}\right) & \text{for } y \leq \mu_y - \theta \sigma_y \\
\frac{1}{\bar{p}\sigma_y \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(y-y_0)^2}{\sigma_y^2}\right) \cdot e^{-\alpha(y-y_0+\theta \sigma_y)-\frac{1}{2}(2y_0+y_0+\theta \sigma_y)^2} & \text{for } y \geq \mu_y - \theta \sigma_y
\end{cases}
\]

Where \(\bar{p}\) is the average probably that someone in the pool will receive a job in the U.S.

After expanding and collecting like terms for the second part of this equation is:

\[
\frac{1}{\bar{p}\sigma_y \sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(y-y_0)^2}{\sigma_y^2}\right) \cdot e^{\frac{(1+y^2)\sigma^2}{(1+y^2)^2}}
\]

Integrating over \(x\) to calculate \(\bar{p}\) yields:

\[
k = \left[ \mu_y^2 - \mu_y \sigma_y (\alpha - \gamma (\mu_y - \theta \sigma_y)) + (\alpha - \gamma (\mu_y - \theta \sigma_y))^2 \right] - \frac{\mu_y^2 - 2 \sigma_y^2 \alpha (\mu_y - \theta \sigma_y) + \frac{\gamma}{2} \sigma_y^2 (\mu_y - \theta \sigma_y)^2}{2 \sigma_y^2}
\]

\(\text{This section demonstrates a proof of the result established by Deaton (2005) using our notation and the proof of the linear case he provides.}\)
\[ \bar{p} = \Phi(-\theta) + \Phi \left( \frac{\theta - \alpha \sigma_y}{(1 + \sigma_y^2)^{1/2}} \right) e^k \]

Substituting this value for \( \bar{p} \) into the second part of (B1) yields

\[ (C6) \quad \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\gamma - \mu_y - \frac{\sigma_y^2 (\alpha + \sigma_y \gamma)}{(1 + \gamma \sigma_y^2)} \right)^2 \left( \frac{\sigma_y^2}{(1 + \gamma \sigma_y^2)} \right)} \cdot \left[ \frac{e^k}{\Phi(-\theta) + \Phi \left( \frac{\theta - \alpha \sigma_y}{(1 + \sigma_y^2)^{1/2}} \right) e^k} \right] \]

Since the share of distribution corresponding to the first part of (B1) decreases with \( \theta \), the overall distribution approaches that of (B6) as \( \theta \) increases. Further, as \( \theta \) becomes large \( \Phi(-\theta) \) approaches 0 while the \( \Phi \left( \frac{\theta - \alpha \sigma_y}{(1 + \sigma_y^2)^{1/2}} \right) \) term approaches 1 allowing the \( e^k \) terms to cancel. This causes (B6) to approach a normal density function with mean \( \mu = \frac{\sigma_y^2 (\alpha + \sigma_y \gamma)}{(1 + \gamma \sigma_y^2)} \) and variance \( \frac{\sigma_y^2}{(1 + \gamma \sigma_y^2)} \). QED
Appendix D – Proof of the Existence of a Unique Interior Solution

By our assumptions above, the country’s objective function is:

\[(D1) \quad U = aqN - \frac{b}{2}N^2\]

Where \(q\) is the measure of pool quality, \(N\) is the number of immigrants admitted and \(a\) and \(b\) are parameters reflecting taste. The policy frontier which constrains this choice was given by equation (22):

\[(D2) \quad E[Y \mid p] = \frac{\bar{Y}}{p} (1 - \Phi(\Phi^{-1}(1 - p) - \sigma_y))\]

Substituting admitted pool average earnings as the quality measure in (D1) and replacing \(p\) with \(N / T\) we obtain:

\[(D3) \quad U = a\frac{\bar{Y}}{N/T} (1 - \Phi(\Phi^{-1}(1 - N / T) - \sigma_y))N - \frac{b}{2}N^2\]

\[(D4) \quad = a\bar{Y}T (1 - \Phi(\Phi^{-1}(1 - N / T) - \sigma_y)) - \frac{b}{2}N^2\]

This problem can be solved by taking the first derivative and setting it equal to 0.

\[(D5) \quad \frac{\partial U}{\partial N} = -a\bar{Y}T\phi(\Phi^{-1}(1 - N / T) - \sigma_y))D_N(\Phi^{-1}(1 - N / T)) - bN\]

Noting, as shown below, that \(D_N(\Phi^{-1}(1 - N / T)) = \frac{-1}{\Phi^{-1}(1 - N / T)}\)

\[
\frac{\partial U}{\partial N} = -a\bar{Y}T\phi(\Phi^{-1}(1 - N / T) - \sigma_y))\frac{-1}{\Phi^{-1}(1 - N / T)} - bN
\]
The optimal immigration policy is found where this function equals 0 or equivalently where:

\[
(D7) \quad \frac{\phi(\Phi^{-1}(1 - N / T) - \sigma_y)}{\phi(\Phi^{-1}(1 - N / T))} = \frac{bN}{aY}
\]

Expanding the left hand side of this function we obtain:

\[
(D8) \quad \frac{\phi(\Phi^{-1}(1 - N / T) - \sigma_y)}{\phi(\Phi^{-1}(1 - N / T))} = \frac{1/2\pi e^{-[(\Phi^{-1}(1-N/T)-\sigma_y)^2]/2}}{1/2\pi e^{-[(\Phi^{-1}(1-N/T))^2]/2}}
\]

\[
= e^{\left[(\Phi^{-1}(1-N/T)-\sigma_y)^2/2\right]-(\Phi^{-1}(1-N/T))^2/2}}
\]

\[
= e^{\left[(\Phi^{-1}(1-N/T)^2+\Phi^{-1}(1-N/T)\sigma_y-\sigma_y^2+\Phi^{-1}(1-N/T)^2)/2\right]}-\sigma_y^2/2\right]
\]

This function is monotonically decreasing in N. Since the Right hand side is monotonically increasing in N, there is exactly one intersection point. Consequently we know that there is a unique solution over the interval N \((0,T)\).

Note: We calculate \(D_N(\Phi^{-1}(1 - N / T))\) as follows:

\[ k = \Phi^{-1}(1 - N / T) \]
\[ \Phi(k) = 1 - N / T \]
\[ \phi(k) \, dk = -(1 / T) \, dN \]

\[ D_N(\Phi^{-1}(1 - N / T)) = \frac{dk}{dN} = \frac{-1}{T\phi(k)} = \frac{-1}{T\phi(\Phi^{-1}(1 - N / T))} \]
References


