Online Appendix for
Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service

Ernesto Dal Bó   Frederico Finan   Martín A. Rossi

February 2013

1 Trustworthiness of our measures

In this section we expand on the discussion contained in the paper substantiating the trustworthiness of our measures.

1.1 Quality

As we discussed in the main body of the paper, our more encompassing measure of quality is the individual's reservation wage proxied by the wage the person received in his latest employment spell. By way of description, in Figure $A2$ in this Appendix we plot the distribution of the applicants' wages in their last employment (dashed line), along with the distribution of wages in 2010 for the population residing in the same municipalities (solid line). The figure also depicts two vertical lines that correspond to the wage offers made in the context of the RDP. As we see from the figure, the distribution of applicants' outside wages is shifted slightly to the right of the distribution for the population, with median wages equal to $3,950$ and $2,571$ Pesos per month respectively. The wage offers of $3,750$ and $5,000$ Pesos correspond approximately to the 65^{th} and 80^{th} percentile of the wage distribution for the population. The shift is compatible with a positive quality selection among the applicants to the RDP.

While measurement error is unavoidable, a bigger concern would be if individuals were systematically misreporting their previous salaries based on the wage offers made by the RDP. In this case, the measurement error would be correlated with treatment. One advantage of our data is that we can study the relationship between the previous wages declared by candidates and the determinants of earnings usually considered in the literature, which are in many cases directly observable. A dilution or distortion of the recovered relationship between declared wages and the determinants of earnings could be expected in the presence of strategically-motivated misreporting. It is therefore pertinent to ask whether our data displays familiar patterns in terms of the determinants of earnings, and whether these patterns change with treatment assignment.

In Table $A2$ of this appendix we present estimates based on variations of a standard Mincerian wage regression, using the wage reported in the last employment as a dependent variable.[1] These applicants' data yield estimates similar to the ones typically found in the existing literature. For instance, the salaries that the men report are on average 15-19 percent higher than those reported by females. This is consistent with estimates within Mexico, but also with those documented in the U.S. literature (e.g., see Blau and Kahn

[1] The results are similar if we use the average of all the reported wages in the last three jobs as the dependent variable.

(2006)). The coefficients on experience and its squared-term imply an experience-earning profile that peaks at around 25 years, and we also find a correlation between wages and the applicant's height; both of which are again consistent with the existing literature (e.g., Persico, Postlewaite, and Silverman (2004)). The returns to schooling for this sample are a familiar 8 percent. We also find, unsurprisingly, that IQ correlates significantly with earnings even after controlling for schooling. The coefficient implies that, compared to the median applicant, applicants who scored 3 points higher were earning salaries that were 6.3 percent higher. If we re-estimate these correlations using data from the 2005 Mexican Family Life Survey (MxFLS) that has been reweighed to match our data, we obtain similar estimates - a result that further validates our applicants' data.[2] In addition to exploring these standard determinants of wages, we also examine whether personal traits predict wages as suggested by the literature in psychology. And again consistent with the literature, the Big 5 traits do correlate with earnings (joint test p-value = 0.08).

Finally, we also tested whether all of these correlations differ between treatment and control and do not find any evidence that they do. Not only are none of the individual correlations statistically significant across treatment and control, but we also fail to reject a test of joint significance (p-value = 0.69). Overall, these results suggest that although wages are self-reported, they do correlate well with the standard predictors of wages, as emphasized in the literature. It seems unlikely that we would have found these results if individuals had been strategically mis-reporting their previous earnings based on the wage announcements of the RDP. For these reasons, we are confident that the applicant's previous wage can serve as an adequate proxy for their reservation wage.

The IQ results are harder to "fake," in that individuals cannot pretend to be smarter than they actually are. But there is some evidence that incentives can increase effort and performance on IQ tests, which would confound our interpretation of the effects of higher wage postings on the intelligence of interested candidates. Importantly, however, these studies find that effort-driven effects tend to be concentrated in the lower third of the distribution (Heckman, Malofeeva, Pinto, and Savelyev 2010). As shown in the paper, we find effects at the upper tail of the distribution. As we emphasized in the paper, subjects in the low wage setting are not without incentives to exert effort. In the high wage condition 62 percent of candidates had a previous wage below the offered (high) wage. In the low wage condition 61 percent of applicants were in the same situation, namely applying for a job paying more than they had earned previously.

[2] The MxFLS is a nationally-representative, longitudinal database which collects a wide range of information on socioeconomic indicators, demographics and health indicators on the Mexican population (Rubalcava and Teruel 2008).

Moreover, in Table $A3$ we examine whether applicants in the higher wage condition are less likely to make careless mistakes in answering a set of individual questions – a pattern of differential mistake rates could reflect differential effort.[3] We do not find any evidence of differences in error rates between treatment and control.

1.2 Public service motivation

Even if applicants did not strategically misreport their wages in view of the verifiability of their report, they may have nonetheless manipulated their responses on questions about pro-social behavior or their inclination for public service. In Table $A4$ of the online appendix, we examine the correlations between PSM and various forms of pro-social behavior. Given that public service motivation is defined in relation to one's inclination to do good for others and shape the well-being of society (Perry and Hondeghem 2008), we would expect PSM to be correlated with one's likelihood to engage in pro-social behavior. This is precisely what we find. Individuals who score higher on the PSM index are more likely to engage in charity, volunteer work, or belong to a political party. They are also more likely to exhibit altruistic tendencies in the hypothetic experimental game. For instance, they are more likely to cooperate in the public goods game and give more to an anonymous player in the dictator game. All these correlations are not only robust to controlling for demographic characteristics (e.g., gender, age, and years of schooling), but also the person's IQ and the Big 5 personality traits, which are strongly correlated with PSM.[4]

Although PSM, and to a lesser extent the Big 5 traits, are good predictors of pro-social behavior, IQ however is not. This raises an important point: If smarter candidates were manipulating their responses to questions on pro-social behavior and other forms of other-regarding attitudes, we would expect a positive association between IQ and these"desirable" traits. But IQ does not correlate positively with any form of pro-social involvement outside of the screening session, and even has a strong negative correlation with altruism, membership with a political party, and the belief that wealth is not important. One may expect strategic misreporting, if it occurs, to be driven by more sophisticated people who to tend to overstate their pro-social and public service motivation. The results in Table $A4$ speak against such possibility, as the correlation between PSM and pro-social behavior is not stronger among individuals with higher IQs.[5]

[3] These questions were selected because they do not require intelligence to be answered and are not likely to be answered in a strategic manner.

[4] The correlation between PSM and IQ is 0.12, whereas the correlation between PSM and Big 5 is 0.54.

[5] The interaction term between Raven's and PSM involves the two variables at their sample means, so as to facilitate the interpretation of the direct effects.

2 Theory

2.1 Proofs for results of the core model presented in the paper

In the main text of the paper we summarize the key takeaways of the selection patterns that obtain in our core model with candidates who differ in terms of quality v and PSM π. We now spell out these results in more detail by considering the different decisions facing a worker in each period and solving the model by backward induction. If offered a job in period 2, a candidate with realized outside opportunity $v+\varepsilon$ will accept the job whenever $v+\varepsilon < w+\pi$, which for a type (v,π) will happen with probability $G(w+\pi-v)$. In period 1, entry decisions depend on the relationship between v and π.

The following proposition, and Figure $A3$ capture the pattern of entry into the applicant pool when candidates differ in terms of quality and PSM.

Proposition 1 *a) v **and** π **independent:** There exists a function $\bar{v}(\pi) = a+\pi$ (with $a>0$, and inverse $\bar{\pi}(v) = -a+v$) describing the locus of all types (v,π) who are indifferent between entering the applicant pool and staying out. Given v, those with $\pi' > \bar{\pi}(v)$ $(< \bar{\pi}(v))$ strictly want to enter (stay out); given π, those with $v' < \bar{v}(\pi)$ $(> \bar{v}(\pi))$ strictly want to enter (stay out).*

*b) v **and** π **positively correlated:** There exists a type $(\bar{v} = m(\bar{\pi}), \bar{\pi})$ who is indifferent between entering the applicant pool and staying out. If $w-c > m(0)$ and $\frac{dm}{d\pi} > 1$, then all types $\pi \leq \bar{\pi}$ enter, and types $\pi > \bar{\pi}$ stay out.*

Proof: a) v **and** π **independent:** The infinite pairs (v,π) who are indifferent between entering and staying out satisfy,

$$v = -c + G(w+\pi-v)(w+\pi) + [1-G(w+\pi-v)][v + E(\varepsilon|\varepsilon > w+\pi-v)]. \quad (1)$$

The implicit function theorem guarantees that a continuous function $\bar{v}(\pi)$ exists, mapping the PSM parameter to the highest quality type $\bar{v}$ that will choose to enter the applicant pool. Rewrite (1) as returns to entry net of expected outside opportunity: $-c + G(w+\pi-v)(w+\pi-v) + \int_{w+\pi-v}^{\infty} \varepsilon g(\varepsilon)\, d\varepsilon = 0$, from which it is clear $\frac{d\bar{v}}{d\pi} = 1$. Note that the assumption $w > c$ yields that if $\pi = 0$, then $\bar{v} > 0$. Thus, the function $\bar{v}(\pi)$ has slope 1 and a positive intercept we label a, yielding $\bar{v} = a+\pi$. Invertibility yields $\bar{\pi} = v - a$. To see the second part of the statement in a), consider the indifferent type (v^i, π^i). The claim is a type (v^i, π') would strictly prefer to enter if $\pi' > \pi^i$. Suppose not. Then net returns to entry must be negative: $-c + G(w+\pi'-v^i)(w+\pi'-v^i) + \int_{w+\pi'-v^i}^{\infty} \varepsilon g(\varepsilon)\, d\varepsilon < 0$. Then, by virtue of

$\pi' > \pi^i$ and net returns being increasing in π (the derivative of the net returns with respect to π is just $G(.) > 0$), we must have $-c + G(w + \pi^i - v^i)(w + \pi^i - v^i) + \int\limits_{w+\pi^i-v^i}^{\infty} \varepsilon g(\varepsilon) d\varepsilon < 0$, which is a contradiction since type π^i was indifferent. Similar logic proves the rest of the statement.

b) v **and** π **positively correlated:** given $v = m(\pi)$, a type π is indifferent between paying the cost c and not iff,

$$m(\pi) = -c + G(w + \pi - m(\pi))(w + \pi) + [1 - G(w + \pi - m(\pi))][m(\pi) + E(\varepsilon|\varepsilon > w + \pi - m(\pi))]. \tag{2}$$

To prove statement b) we need the LHS of (2) to be smaller (larger) than the RHS for $\pi \leq \bar{\pi}$ ($> \bar{\pi}$). A sufficient condition for this is that the LHS have a smaller intercept and a steeper slope than the RHS (in the $(\pi, m(\pi) = v)$ space). For lower intercept we need,

$$m(0) < -c + G(w - m(0))w + [1 - G(w - m(0))][m(0) + E(\varepsilon|\varepsilon > w - m(0))],$$

which obtains whenever $w - c > m(0)$. The LHS has a steeper slope than the RHS iff $\frac{dm}{d\pi} > G(.) + [1 - G(.)]\frac{dm}{d\pi}$ or $\frac{dm}{d\pi} > 1$ as assumed in the main text.■

This proposition tells us that the type space (v, π) can be divided into two sets containing respectively the types who apply and who do not. When the type dimensions are independent, each candidate dimension is selected in opposite directions: the relatively low quality but relatively high PSM individuals opt in. The reason is that the expected value of entering the pool increases in PSM but decreases in quality. Thus, if v and π are independent in the population, any positive correlation among applicants is due to self-selection, as high quality types will only apply if they also have a very high PSM. When the type dimensions are positively correlated along the function $m(\pi)$ with $m' > 1$, the applicant pool contains all types (v, π) up to $(\bar{v} = m(\bar{\pi}), \bar{\pi})$; in other words, the relatively low types apply, and the relatively high types stay out of the pool. The intuition is that when the two dimensions are collapsed into one, the opposite selection forces of the independent case net out in one direction or another depending on the slope of $m(\pi)$. If quality rises with PSM more than one for one, the quality effect dominates and the relatively high types stay out. (In this case, the applicant pool will display a positive correlation between quality and PSM that is inherited from the population.)

This part of our analysis is related to an interesting model by (Delfgaauw and Dur 2007). The authors also derive an indifference condition in the space of quality and PSM in the context of a competitive economy, but abstract from the application costs, and the option value of paying them. This difference plays a role in part b) of proposition 1, and in the contrasting effects of higher wages on the applicant pool that we characterize in Proposition 1 in the paper, which we state and prove here.

Proposition 1 (in the paper's main text):

a) Given the assumptions of our model, an increase in wages increases the size and average quality of the applicant pool.

b) In the case when PSM and quality are independent in the population, an increase in wages decreases the average PSM of the applicant pool.

c) In the case when PSM and quality are positively correlated according to the function $m(\pi)$*, an increase in wages increases the average PSM of the applicant pool.*

Proof: a) In the independent case the function mapping π to the highest quality type who applies is characterized implicitly by (1), and it depends on w, so we write $\bar{v}(\pi, w) = a(w)+\pi$. Showing this function increases in w will establish statement a) for the independent case. By invertibility $\bar{\pi}(v, w) = v - a(w)$, so the same proof will establish statement b). Rewriting (1) in the implicit function, differentiating wrt w and canceling terms, we get $\frac{d\bar{v}}{dw} = 1 > 0$, which proves a) and b) for the independent case.

In the positively correlated case, we can simultaneously prove the statements in a) and c) by showing that the indifferent type $\bar{\pi}$ increases in w yielding that $\bar{v} = m(\bar{\pi})$ must also increase by virtue of m being increasing (recall $\bar{\pi}$ is the locus where $m(\pi)$ cuts the RHS of (2) from below). Again, it is easily seen the implicit function $\bar{\pi}(w)$ exists as characterized by (2). Writing this expression in the implicit function, differentiating with respect to w and rearranging, we get,

$$\frac{dm}{d\bar{\pi}}\frac{d\bar{\pi}}{dw} = G(.) + \frac{d\bar{\pi}}{dw}\left\{G + \frac{dm}{d\pi}(1-G)\right\} \tag{3}$$

$$\frac{d\bar{\pi}}{dw} = \frac{1}{\frac{dm}{d\bar{\pi}} - 1} > 0 \tag{4}$$

where the inequality follows from the assumption $\frac{dm}{d\pi} > 1$.■

These comparative statics are quite intuitive from inspection of Figure $A3$–the effect of an increase in wages is to lower the function $\pi(v) = v - a(w)$.

2.2 Model with job attributes

We expand the model to study the effects of job characteristics on acceptance, and the role of wages in overcoming the problems posed by undesirable job characteristics. Examples of these characteristics are geographic location, and a safe or affluent social environment. To do this, we take the basic version of the model with $\rho = 1$, utility from the public sector job $u(w, X)$ and a single interview cost $c > 0$. The job offer, when it materializes, will be associated with a wage $w = \{\underline{w}, \bar{w}\}$ known to the candidate at the time of application, and a job type parameter x, which is the only element in X. To make things as simple as possible we take the attribute x to affect linearly the utility from taking the job, so $u(w, x) = w + x$. The job characteristic x is only revealed to the candidate at the time when the job offer is made, and is drawn from a cdf $H(.)$ with associated density $h(.)$ with unbounded support. The realization of x is independent from that of ε.

In **period 2**, and given an applicant pool with highest quality $\bar{v}$, an individual of type v and realized shock ε accepts the job whenever $v + \varepsilon \leq w + x$. That is, if $\varepsilon - x \leq w - v$. Note that acceptance now does not just depend on the realization of ε, but also of x, or on the realization of a new random variable $z = \varepsilon - x$. Define the cumulative distribution function (cdf) $J_z(w - v) \equiv (z = \varepsilon - x \leq w - v)$ with associated density $j_z(w - v)$. The cdf is given by the usual convolution formula, yielding,

$$
\begin{aligned}
J_z(w - v) &\equiv P(z \leq w - v) = \int_{-\infty}^{\infty} P(\varepsilon < w - v + x)\, h(x)\, dx \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{w-v+x} g(\varepsilon)\, d\varepsilon \right) h(x)\, dx = \int_{-\infty}^{\infty} G(w - v + x)\, h(x)\, dx,
\end{aligned}
$$

while the associated density is $j_z(w - v) \equiv P(dz = w - v) = \int_{-\infty}^{\infty} g(w - v + x)\, h(x)\, dx$.

In **period 1** an individual of type v applies if and only if the usual entry condition $v \leq \psi(v)$ holds, where this condition now reads,

$$
v \leq -c + J_z(w - v)\{w + E[x|z \leq w - v]\} + [1 - J_z(w - v)]\{v + E[\varepsilon|z > w - v]\}. \quad (5)
$$

We can now state,

Lemma 1 *If the application cost c is small relative to the utility from the wage and expected job attributes, the entry equilibrium is analogous to that in the basic model: there exists a finite type $\bar{v}$ that is indifferent between attending the interview and staying out. All $v \leq \bar{v}$ prefer to attend and enter the candidate pool, while all $v > \bar{v}$ stay out, and the separating type $\bar{v}$ is increasing in w.*

Proof: Positive, bounded entry is ensured by $\psi(0) > 0$ and $\psi' < 1$, so we show these inequalities hold under the assumptions made.

$\psi(0) > 0$: Tedious algebra shows that the conditional expectations in the entry condition can be written as,

$$
\begin{aligned}
E\left(x|z\leq w-v\right) &= \int\limits_{-\infty}^{\infty} x.h\left(x|z\leq w-v\right)dx \\
&= \frac{1}{J\left(w-v\right)}\int\limits_{-\infty}^{\infty} x.\left(\int_{-\infty}^{w-v} g\left(z+x\right)dz\right)h\left(x\right)dx. \\
E\left[\varepsilon|z\geq w-v\right] &= \int_{-\infty}^{\infty}\varepsilon g\left(\varepsilon|z\geq w-v\right)d\varepsilon \\
&= \frac{1}{1-J\left(w-v\right)}\int_{-\infty}^{\infty}\varepsilon.\left(\int_{w-v}^{\infty} h\left(\varepsilon-z\right)dz\right)g\left(\varepsilon\right)d\varepsilon.
\end{aligned}
$$

Then we can rewrite $\psi\left(v\right)$ in the entry condition as,

$$
\begin{aligned}
\psi &= -c+J_z\left(w-v\right)w+\int\limits_{-\infty}^{\infty} x.\left(\int_{-\infty}^{w-v} g\left(z+x\right)dz\right)h\left(x\right)dx+ \\
&\quad +\left[1-J_z\left(w-v\right)\right]v+\int\limits_{-\infty}^{\infty}\varepsilon.\left(\int_{w-v}^{\infty} h\left(\varepsilon-z\right)dz\right)g\left(\varepsilon\right)d\varepsilon.
\end{aligned}
$$

The value of this expression at $v=0$ is,

$$
\begin{aligned}
&-c+J_z\left(w\right)\left[w+\frac{1}{J\left(w\right)}\int\limits_{-\infty}^{\infty} x.\left(\int_{-\infty}^{w} g\left(z+x\right)dz\right)h\left(x\right)dx\right]+ \\
&+\int\limits_{-\infty}^{\infty}\varepsilon.\left(\int_{w-v}^{\infty} h\left(\varepsilon-z\right)dz\right)g\left(\varepsilon\right)d\varepsilon.
\end{aligned}
$$

The gross value of applying (given by the second and third terms in the latter expression) increases in w, so w high enough guarantees $\psi\left(0\right)>0$ and a positive measure of applicants.

$\psi'<1$: Note that,

$$
\frac{d\psi\left(v\right)}{dv}=-j\left(.\right)w-\int\limits_{-\infty}^{\infty} x.g\left(w-v+x\right)h\left(x\right)dx+j\left(.\right)v+1-J\left(.\right)+\int\limits_{-\infty}^{\infty}\varepsilon.h\left(\varepsilon-w+v\right)g\left(\varepsilon\right)d\varepsilon.
$$

Making the change of variable $x = \varepsilon - w + v$, the last equality becomes,

$$\begin{aligned}\frac{d\psi(v)}{dv} &= 1 - J(.) - j(.)(w-v) + (w-v)\int_{-\infty}^{\infty} g(\varepsilon) h(\varepsilon - w + v)\, dx - \\ &\quad - \int_{-\infty}^{\infty} \varepsilon . g(\varepsilon) h(\varepsilon - w + v)\, dx + \int_{-\infty}^{\infty} \varepsilon . h(\varepsilon - w + v) g(\varepsilon)\, d\varepsilon,\end{aligned}$$

and given $(w-v)\int_{-\infty}^{\infty} g(\varepsilon) h(\varepsilon - w + v)\, dx = j(.)$, we get,

$$\frac{d\psi(v)}{dv} = 1 - J(.) < 1,$$

implying there exists a finite, highest type $\bar{v}$ who applies.

To see the marginal type $\bar{v}$ increases in w, note that the applicant pool is determined by the equality $\bar{v}(w) = \psi(\bar{v}(w))$. Totally differentiating,

$$\frac{d\bar{v}}{dw} = \frac{d\psi(v)}{dv}\frac{d\bar{v}}{dw} + \frac{d\psi}{dw},$$

so $\frac{d\bar{v}}{dw}\left(1 - \frac{d\psi(v)}{dv}\right) = \frac{d\psi}{dw}$. It is easy to show that $\frac{d\psi(v)}{dv} = 1 - \frac{d\psi}{dw}$, readily implying $\frac{d\bar{v}}{dw} = 1 > 0$.■

The last lemma establishes that the added complexity of uncertainty over job attributes does not alter the essence of self-selection into the applicant pool. As in the basic model, a wage w leads to an applicant pool of size $F(\bar{v}(w))$. Therefore, conditional on the wage and the realized attribute x, the acceptance rate in period 2 is

$$\begin{aligned}P(acc|w,x) &= \int_0^{\bar{v}(w)} P(\varepsilon < w - v + x)\frac{f(v)}{F(\bar{v}(w))}dv \\ &= \int_0^{\bar{v}(w)} G(w - v + x)\frac{f(v)}{F(\bar{v}(w))}dv\end{aligned}$$

Now we can establish,

Proposition 2 *a) The conditional acceptance rate $P(acc|w,x)$ is increasing in x.*

b) The interaction of wages and job attribute x is generally ambiguous. However, if the densities g and f are decreasing, the interaction is negative.

Proof: a) Note $\frac{dP(acc|w,x)}{dx} = \int_0^{\bar{v}(w)} g(w - v + x)\frac{f(v)}{F(\bar{v}(w))}dv > 0$.

b) Note

$$\frac{d^2P(acc|w,x)}{dxdw} = \frac{f(\bar{v}(w))}{F(\bar{v}(w))}\left[g(w-\bar{v}(w)+x) - \int_0^{\bar{v}(w)} g(w-v+x)\frac{f(v)}{F(\bar{v}(w))}dv\right]\frac{d\bar{v}}{dw} + \\ + \int_0^{\bar{v}(w)} \frac{dg(w-v+x)}{dw}\frac{f(v)}{F(\bar{v}(w))}dv.$$

Now consider g increasing (decreasing). Then the term in square brackets is negative (positive), while $\int_0^{\bar{v}(w)} \frac{dg(w-v+x)}{dw}\frac{f(v)}{F(\bar{v}(w))}dv > 0 (< 0)$, and the sign of $\frac{d^2P(acc|w,x)}{dxdw}$ is ambiguous. The term in square brackets reflects changes in the applicant pool driven by self-selection, while the integral $\int_0^{\bar{v}(w)} \frac{dg(w-v+x)}{dw}\frac{f(v)}{F(\bar{v}(w))}dv$ reflects direct effects of wages on acceptance rates by inframarginal types.

To verify that $g' < 0$ and $f' < 0$ yield a negative interaction, assume $g' < 0, f' < 0$ and recall $\frac{dg(w-v+x)}{dw} = -\frac{dg(w-v+x)}{dv}$, and that, from lemma **1**), $\frac{d\bar{v}}{dw} = 1$. Then the interaction is negative iff

$$f(\bar{v}(w))\left[g(w-\bar{v}(w)+x) - \int_0^{\bar{v}(w)} g(w-v+x)\frac{f(v)}{F(\bar{v}(w))}dv\right] < \int_0^{\bar{v}(w)} \frac{dg(w-v+x)}{dv} f(v)\,dv.$$

Iterated integration by parts and some rearranging yield,

$$\frac{\int_0^{\bar{v}(w)} f(v)\frac{dg}{dv}dv}{\int_0^{\bar{v}(w)} F(v)\frac{dg}{dv}dv} > \frac{f(\bar{v}(w))}{F(\bar{v}(w))}.$$

By virtue of $g' < 0$, $\frac{dg}{dv} > 0$, and we can write,

$$\frac{\int_0^{\bar{v}} f(v)\frac{\frac{dg}{dv}}{g(\bar{v})-g(0)}dv}{\int_0^{\bar{v}} F(v)\frac{\frac{dg}{dv}}{g(\bar{v})-g(0)}dv} > \frac{f(\bar{v})}{F(\bar{v})}.$$

Note that $\int_0^{\bar{v}} \frac{\frac{dg}{dv}}{g(\bar{v})-g(0)}dv = 1$, and therefore for a generic function $\phi(v)$ we have $\int_0^{\bar{v}} \phi(v)\frac{\frac{dg}{dv}}{g(\bar{v})-g(0)}dv > (<) \int_0^{\bar{v}} \phi(\bar{v})\frac{\frac{dg}{dv}}{g(\bar{v})-g(0)}dv \equiv \phi(\bar{v})$ whenever $\phi(v)$ is decreasing (increasing), yielding the inequality.■

This proposition tells us that acceptance rates improve with the realization of good job attributes, and that if high outside options and high quality types are less likely than low ones, then the interaction of attributes and wages is negative. This means higher attributes make less of a difference when wages are high, or alternatively, that wages are particularly useful at mitigating the negative effects of bad job attributes.

2.3 Model with non-uniform probability of getting an offer

2.3.1 Probability depends on own quality

As in our basic model, we assume the public sector job yields a utility u, and that the value that a type v will get in a private sector position is $v + \varepsilon$. But now suppose candidates believe not everyone will get an offer, and that the chance of getting an offer increases in one's type. For example, the candidates could believe the test used by the RDP would reveal each candidate's v type with noise, and that the program would offer a job to everyone who performed above a given mark. Suppose the test on individual i yields a score $s_i = v_i + \mu$, with μ having zero expectation, and distributed with CDF H, and associated density h. The employer then makes an offer to anyone for whom $s_i > t$, where t captures the employability "threshold." Then the probability of being made an offer is $\rho(v) = 1 - H(t - v_i)$ which is increasing in v_i. We drop subscripts in what follows for simplicity. We impose the assumption $u > \frac{c}{\rho(0)}$, which guarantees that the lowest type still perceives it as worthwhile to show up to the exam, and a second assumption that the elasticity of the probability ρ is not too high. Namely, we assume $\eta_\rho \equiv \frac{\rho'(v)}{\rho(v)} v < \frac{v}{[u-v-E(\varepsilon|\varepsilon<u-v)]}$. Clearly, this assumption entails a condition on the distribution of noise H, namely that $\frac{h(t-v)}{1-H(t-v)} < \frac{1}{[u-v-E(\varepsilon|\varepsilon<u-v)]}$, or that the hazard rate of the noise in the test not be too high. The hazard rate reflects the marginal benefit of having a higher type in terms of a higher chance of scoring above the employability threshold. The bound is given by the inverse of $u - v - E(\varepsilon|\varepsilon < u - v)$ which reflects the option value of attending the exam: i.e., the expected extra utility the public sector job will yield in the event in which the candidate is unlucky with the realization of her reservation utility.

The timing of interaction is as in the main model:

Period 2. Each individual that paid c and attended the interview receives an offer with probability $\rho(v)$. If an offer is received, the individual accepts it if and only if $u \geq v + \varepsilon$, (ε and μ are independent) which occurs with probability $G(u - v)$. Thus, individuals of higher quality are less likely to accept the job. Individuals that reject the offer, do not receive one, or that did not attend the interview, get $v + \varepsilon$.

Period 1. The candidate must decide whether to pay c to buy the option embodied in a (potential) offer that might be better than his realized reservation utility. The worker will decide to pay the cost c if $v \leq \psi(v, u, \rho)$, where $\psi(v, u, \rho)$ is the expected payoff from paying and attending the interview, and given by the expression,

$$\begin{aligned} \psi &= -c + G(u-v)\{\rho(v)u + (1-\rho(v))[v + E(\varepsilon|\varepsilon < u - v)]\} \\ &+ [1 - G(u-v)](v + E(\varepsilon|\varepsilon > u - v)). \end{aligned} \tag{6}$$

The value of attending the interview is that if $\varepsilon < u - v$ (realized reservation utility is low), then the candidate has, with probability $\rho(v)$, the option to take the job. We can now establish,

Proposition 3 *If* $\eta_\rho < \frac{v}{[u-v-E(\varepsilon|\varepsilon<u-v)]}$, *then*

a) There exists a finite type $\bar{v}$ that is indifferent between attending the interview or not. All $v \leq \bar{v}$ prefer to attend and enter the candidate pool, while all $v > \bar{v}$ stay out.

b) The separating type $\bar{v}$ is increasing in the value of the job $u = U(w)$.

Proof: a) It follows from the fact that the function ψ crosses the function v from above once in the space (v, ψ). To see this, note ψ has intercept

$$\begin{aligned}
\psi(0) &= -c + G(u)\{\rho(0)u + (1-\rho(0))[E(\varepsilon|\varepsilon < u)]\} + [1 - G(u)](E(\varepsilon|\varepsilon > u)) \\
&= -c + G(u)\rho(0)u + (1-\rho(0))G(u)[E(\varepsilon|\varepsilon < u)] + [1 - G(u)](E(\varepsilon|\varepsilon > u)) \\
&= -c + \rho(0)\{G(u)u + [1 - G(u)](E(\varepsilon|\varepsilon > u))\}.
\end{aligned}$$

Note this intercept is positive $-c + \rho(0)\{G(u)u + [1-G(u)]E(\varepsilon|\varepsilon > u)\} > 0$ by virtue of assumption $u > \frac{c}{\rho(0)}$. The second step is to show that the function $\psi(v)$ has slope less than one. To see this, let us rewrite the function ψ as

$$\begin{aligned}
\psi &= -c + G(u-v)\left\{\rho(v)u + (1-\rho(v))\left[v + \int_{-\infty}^{u-v} \frac{\varepsilon g(\varepsilon)}{G(u-v)} d\varepsilon\right]\right\} \\
&+ [1 - G(u-v)]\left(v + \int_{u-v}^{\infty} \frac{\varepsilon g(\varepsilon)}{1 - G(u-v)} d\varepsilon\right).
\end{aligned} \tag{7}$$

$$\begin{aligned}
\psi &= -c + G(u-v)\{\rho(v)u + (1-\rho(v))v\} + (1-\rho(v))\int_{-\infty}^{u-v} \varepsilon g(\varepsilon)\, d\varepsilon \\
&+ [1 - G(u-v)]v + \int_{u-v}^{\infty} \varepsilon g(\varepsilon)\, d\varepsilon.
\end{aligned} \tag{8}$$

The slope of $\psi(v)$ can easily be computed to,

$$\frac{d\psi}{dv} = 1 - \rho(v)G(u-v) + \rho'(v)G(u-v)[u - v - E(\varepsilon|\varepsilon < u-v)].$$

Thus, $\frac{d\psi}{dv} < 1$ whenever,

$$\begin{aligned} \rho(v) - \rho'(v)\left[u - v - E(\varepsilon|\varepsilon < u - v)\right] &> 0 \\ \rho(v) &> \rho'(v)\left[u - v - E(\varepsilon|\varepsilon < u - v)\right] \\ \frac{v}{\left[u - v - E(\varepsilon|\varepsilon < u - v)\right]} &> \frac{\rho'(v)}{\rho(v)} v \equiv \eta_\rho. \end{aligned}$$

which holds by virtue of the condition invoked in the proposition.

b) We need to show that ψ is increasing in u. Note $\frac{d\psi}{du} = \rho(v) G(u - v) > 0$.■

This proposition offers a simple characterization. Under the assumption $\rho(0) u > c$ we will have a non empty applicant pool (this is the analog of the assumption $u > c$ in our basic model, requiring that the direct expected return of showing up for an interview compensates the cost for the lowest type). In the basic model where $\rho(v)$ adopts a uniform value ρ, high enough types do not enter as they are unlikely to exercise the option of taking a public sector job. When $\rho(v)$ is increasing, the value of entering the applicant pool reflects two forces. Higher types are still less likely to want to exercise the option to take the public sector job, but they are more likely to do well in the exam and to be made an offer, and this contributes to their possibly having a higher interest in applying. Under the condition that $\eta_\rho > \frac{v}{[u-v-E(\varepsilon|\varepsilon<u-v)]}$, this second force is weak enough that the selection pattern, and comparative statics of u, characterized in the basic model are preserved.

Different selection patterns It is worth analyzing a few other cases of interest. The first is one where the condition $\eta_\rho < \frac{v}{[u-v-E(\varepsilon|\varepsilon<u-v)]}$ does not hold. Selection patterns could differ as there is no longer a guarantee that the function ψ intersects v, or that if an intersection exists it is unique. The two following cases could obtain:

1. **All in** (no intersection): If $\psi(v) > v$ for all v, then everyone enters the pool. This is possible conditional on wages w being high enough. Judging from the data, this case did not seem to obtain in the baseline with $w = \$3,750$ - the increase in wages tended to increase the number of candidates.

2. **Non-monotone selection** (multiple intersections): under the assumption $u > \frac{c}{\rho(0)}$, low enough types will always apply, but if the slope of $\psi(v)$ can be larger than 1, we could have multiple intersections of $\psi(v)$ and v. So the case can obtain where there are two intersections v' and v'', so that both low types ($v < v'$) and high types ($v > v''$) enter, while intermediate types ($v' \leq v \leq v''$) stay out (as $\psi(v)$ crosses v from above at v' and from below at v''). The high types come in because the higher chances of getting an offer justify their paying the cost c of sitting the exam. In this non-monotone selection scenario, an increase in

w, raising ψ, would increase v' and decrease v'', and as a result it would raise the size of the applicant pool, but it would have ambiguous effects on the average quality. As v' increases, the average quality of the low types who apply would be higher, but given the decrease in v'', the average quality of the high types who apply would be lower. With three or more intersections we could of course have various different patterns of alternating entry/no entry decisions for different values of v.

2.3.2 Candidates believe their chances of employment depend not only on their own type but also on that of other candidates

In a situation in which the employer has scarce slots that will be awarded to those perceived to be the best candidates, rational individuals may consider that their own chances of getting a job depend not just on their own type but also that of the other applicants. We saw that if candidates believe a higher wage entails a higher employability threshold, a discouragement effect could set in. The possibility that higher wages may attract more competition by high types is an additional reason for such discouragement. A note of realism suggests caution– the discouragement may be unlikely if the lower types lack the cognitive sophistication to predict how an increase in competition from the higher types translates into a lower chance of appointment. So there might be a contradiction in terms to working out how low quality individuals would behave in a model in which they are also assumed to be fully rational. But in the interest of tracing out how fully rational candidates would make their decisions, we focus here on two different scenarios. The objective is to show that the results in our basic model can be preserved, but that as in the case in the previous section, this is not a general property. In performing our analysis here we have benefitted from reading the interesting work by Morgan, Sisak, and Vardy (2012), who rigorously analyze a more general setting involving general equilibrium effects.

Exam is perfectly informative To show how our basic results could be upset if candidates make their entry decisions with an eye to the behavior of others, consider first a situation characterized by the following three conditions: first, the exam reveals the candidate's type without noise (or, rather that this is what candidates believe), second, the employer wants to appoint the highest possible types, and third, there is a small set of positions available so a candidate will receive an offer iff having the highest type. Consider for convenience that the type space is bounded above, so types are distributed according to the function $F(v)$ with associated density $f(v)$ over the support $[0, \hat{v}]$. Now define as $v' \leq \hat{v}$ as the highest type that would consider entry worthwhile under the prospect of getting an offer with probability one. Define also a selection pattern as a set $V \subset [0, \hat{v}]$ comprising the types who decide to enter. Let us index each possible set of entrants V with the highest type in that set, so we write V_v if the highest entrant in the set V is v. Note that given the exam will perfectly reveal each applicant's type, any set of entrants V_v with $v < v'$ can be beaten by a type v' entering, and that this type would desire to enter. Thus, any type $v < v'$ will decide to stay out, as it does not pay to incur the cost c to have zero chance of getting an offer. In this modified model, the selection pattern characterized in our basic treatment unravels, and only the highest type applies who finds it worthwhile to do so given a certain future job offer. That is exactly the type $\bar{v}$ that characterized the upper bound of the applicant pool in our basic model. Thus, increases in wages in the modified model will have similar effects on the highest type who applies as in our basic model, but in this modified model it is no longer true that the applicant pool increases in size with higher wages (except when the density f is increasing around v').

Exam is imperfectly informative We return to the noisy exam technology used earlier, where a type v obtains a score $s = v + \mu$, with μ drawn from a mean zero cdf $H(\mu)$ with associated density h. Again we assume there exists a highest type $\hat{v}$. The candidates could have entertained the notion that only the top $T\%$ of exam performers would be offered a job. In this scenario it is possible to obtain selection patterns that differ from that in our basic model, in similar fashion to the case in which candidates expected offers to take place iff their score s exceeded a given threshold. Here we focus attention on showing that there exists a set of conditions under which the results in the basic model are preserved.

Consider the family of selection patterns $V_{\bar{v}}$ that have as entrants all types v below a given type $\bar{v} \leq \hat{v}$. Given that the exam reveals the taker's type with noise, it is clear that a set of entrants $V_{\bar{v}}$ generates a distribution of scores $S(s|\bar{v})$, and that an applicant pool with a higher marginal entrant generates a higher distribution of scores (i.e., $S(s|\bar{v}) < S(s|\bar{v}')$ iff $\bar{v}' < \bar{v}$. Now consider a putative equilibrium in which all types up to $\bar{v}$ enter. Given the top $T\%$ scorers will obtain an offer, the score that separates offer recipients from rejects is a threshold t satisfying $T = 1 - S(t|\bar{v})$, or equivalently, $t = S^{-1}(1 - T|\bar{v})$, where the threshold obviously increases with the index $\bar{v}$ of the quality of the pool. Now any given type v knows that her chance of meeting or exceeding that threshold, and her probability of getting an offer, is $\rho(v, \bar{v}) = 1 - H(t - v) = 1 - H(S^{-1}(1 - T|\bar{v}) - v)$.

As before, we can write the entry condition as the net return of sitting the exam $\psi(v, \bar{v}) - v \geq 0$, although the function ψ now depends on the pool of entrants $\bar{v}$, as this affects the exam threshold. Written in full, this condition reads just like that in (6) except that $\rho(v)$ is now $\rho(v, \bar{v}) = 1 - H[S^{-1}(1 - T|\bar{v}) - v]$, which increases with more vacancies T, and which decreases with a higher quality pool $\bar{v}$ showing up.

Now we impose the following three assumptions:

1. The lowest type $v = 0$ prefers to enter even if all types enter and the chance of an offer is at a minimum:

$$-c + G(u) \left\{ \begin{array}{c} \left[1 - H\left[S^{-1}(1 - T|\hat{v})\right]\right] w \\ +H\left[S^{-1}(1 - T|\hat{v})\right](E(\varepsilon|\varepsilon < u)) \end{array} \right\} + [1 - G(u)] E[\varepsilon|\varepsilon > u] > 0.$$

2. The highest type $v = \bar{v}$ prefers to stay out even if the applicant pool is of minimal quality $\bar{v} = 0$, and the chances of an offer are maximal:

$$-c+G(u-\hat{v}) \left\{ \begin{array}{c} \left[1 - H\left[S^{-1}(1 - T|0)\right]\right] w \\ +H\left[S^{-1}(1 - T|0)\right](\hat{v} + E(\varepsilon|\varepsilon < u - \hat{v})) \end{array} \right\} + [1 - G(u - \hat{v})](\hat{v} + E[\varepsilon|\varepsilon > u - \hat{v}]) < \hat{v}.$$

3. The hazard rate $\frac{h(t-v)}{1-H(t-v)}$ is not too high (i.e., it satisfies the condition $\frac{h(t-v)}{1-H(t-v)} < \frac{1}{[u-v-E(\varepsilon|\varepsilon<u-v)]}$ invoked in section 2.3.1)

The third condition was discussed earlier. The first two conditions are more likely to be met when the distance $\hat{v}$ between the lowest and highest type is small relative to the amount of noise in the exam, so that the chances of an offer for a low type are not too low even if the highest types apply, and if the chances of an offer to a high type are not too high even if only the lowest types apply.

We can now establish a result characterizing a selection pattern analogous to that in our basic model,

Proposition 4 *There exists a unique entry equilibrium in which all types $v \leq v^* \in (0, \hat{v})$ enter, and all types $v > v^*$ stay out. The marginal entrant v^*increases with the wage w.*

A sketch of the proof for this result is as follows. Note that both $\psi(v, \bar{v})$ and v are continuous. Fixing $\bar{v}$, assumptions 1-3 guarantee the existence of a unique type $\bar{v}'$ who is indifferent between entering and not. In other words, the entry indifference condition

$$\psi(v, \bar{v}) - v = 0$$

constitutes a (continuous) map $\bar{v}' = \Psi(\bar{v}) : [0, \hat{v}] \rightarrow [0, \hat{v}]$. For a given applicant pool $\bar{v}$, Ψ yields another set of entrants $\bar{v}'$. An equilibrium requires a fixed point $\bar{v} = \Psi(\bar{v})$. Given continuity, a sufficient condition for the existence of a unique fixed point is that Ψ be decreasing. And this is indeed the case from the fact that ψ is increasing in $\rho(v, \bar{v})$ but $\rho(v, \bar{v})$ is decreasing in $\bar{v}$. That v^* must increase with w follows from the fact that $\psi(v, \bar{v})$ increases in w, and then so does $\Psi(\bar{v})$, forcing its intersection with $\bar{v}$ to increase.

2.4 Candidates believe that a higher wage signals job characteristics

Candidates might infer different things from a higher announced wage. Here we analyze two possibilities.

2.4.1 Higher wages as a signal of job satisfaction

Individuals might infer that a job carrying a higher wage might entail more sophisticated and interesting tasks. In this case, the utility conditional on taking the job would be $u = w + i(w)$, where $i(w)$ is an intrinsic taste for the job that is increasing in the announced wage. This variation is equivalent to an even higher wage announcement, and from our basic model it is directly predicted to increase the size and average quality of the pool of applicants. Conversely, if candidates inferred that a higher wage must be there to compensate for unsavory or difficult tasks, the function $i(w)$ would be decreasing. Then, the effect of wages we analyze in the basic model would be mitigated, and, depending on the strength of the effects through intrinsic motivation, could be reversed.

2.4.2 Higher wages as a signal of a higher threshold for employability

Suppose that candidates believe the threshold t to increase with the wage announced: $t = t(w)$, with $\frac{dt}{dw} > 0$. Thus, a job announcement carrying a higher wage induces everyone to perceive lower chances of getting employed: $\frac{d\rho}{dw} = -h(t-v)\frac{dt}{dw} < 0$, so the value of attending the interview is lower. Thus, the positive selection effects of a higher wage could be mitigated.

To see this, note that we can now write $\psi(\rho(v,w), w)$, where $\frac{d\rho}{dw} < 0$. Thus, the total differential $\frac{d\psi}{dw}$ is,

$$\frac{d\psi}{dw} = \rho(v)\,G(w-v) + G(.)\left(w - v - \int_{-\infty}^{w-v} \varepsilon g(\varepsilon)\, d\varepsilon\right)\frac{d\rho}{dw},$$

where the second term is negative. If the discouragement effect is sufficiently strong, then higher wages could lead to a downward shift of the function ψ, and therefore to negative selection.

References

Blau, F. D. and L. M. Kahn (2006, October). The U.S. gender pay gap in the 1990s: Slowing convergence. *Industrial and Labor Relations Review 60*(1), 45–66.

Delfgaauw, J. and R. Dur (2007, April). Signaling and screening of workers' motivation. *Journal of Economic Behavior & Organization 62*(4), 605–624.

Heckman, J. J., L. Malofeeva, R. Pinto, and P. A. Savelyev (2010). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *Unpublished*.

Morgan, J., D. Sisak, and F. Vardy (2012). The merits of meritocracy. *Unpublished*.

Perry, J. and A. Hondeghem (2008). *Motivation in public management: the call of public service*. Oxford University Press.

Persico, N., A. Postlewaite, and D. Silverman (2004, October). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy 112*(5), 1019–1053.

Rubalcava, L. and G. Teruel (2008). User's guide for the Mexican family life survey second wave. *MxFLS Documentation*.

Vivir Mejor

ANUNCIO PARA EMPLEO

ESTADOS UNIDOS MEXICANOS

GOBIERNO FEDERAL

El Gobierno Federal, a través del Proyecto para el Desarrollo de Regiones Vulnerables, va a contratar

PROMOTORES SOCIALES

REQUISITOS:

1. SER MEXICANOS (HOMBRES Y MUJERES)
2. MAYORES DE 18 AÑOS
3. DEDICACIÓN DE TIEMPO COMPLETO
4. CON DISPOSICION PARA TRABAJAR Y VIVIR EN COMUNIDADES DE ALTA Y MUY ALTA MARGINACIÓN
5. CAPACES DE EXPRESARSE CLARAMENTE POR ESCRITO Y VERBALMENTE
6. CON BUEN TRATO INTERPERSONAL

RESPONSABILIDADES

El Promotor Social trabajará directamente con las autoridades y los habitantes de municipios de alta y muy alta marginación, proporcionando apoyo para la identificación de necesidades y demandas sociales de la comunidad, facilitando procesos de organización, orientando a la población y sus autoridades sobre las posibilidades de apoyo del gobierno federal para el desarrollo de sus comunidades, y facilitando su vinculación.

$$ SALARIO ATRACTIVO $$

Para registrarte en el proceso de selección y obtener mayores informes sobre el trabajo, llama al

01-800-XXX-XXXX

(teléfono gratuito; no se cobra la llamada)

O escribe al correo electrónico*: XXXX@gmail.com

*Si usted decide escribir un correo electrónico, favor de mencionar su nombre completo, así como el nombre de la localidad en donde usted vio este anuncio.

FECHA LÍMITE PARA REGISTRARTE: Viernes, 4 de junio, 2011, 17 horas

FIGURE A1: THE JOB ANNOUNCEMENT

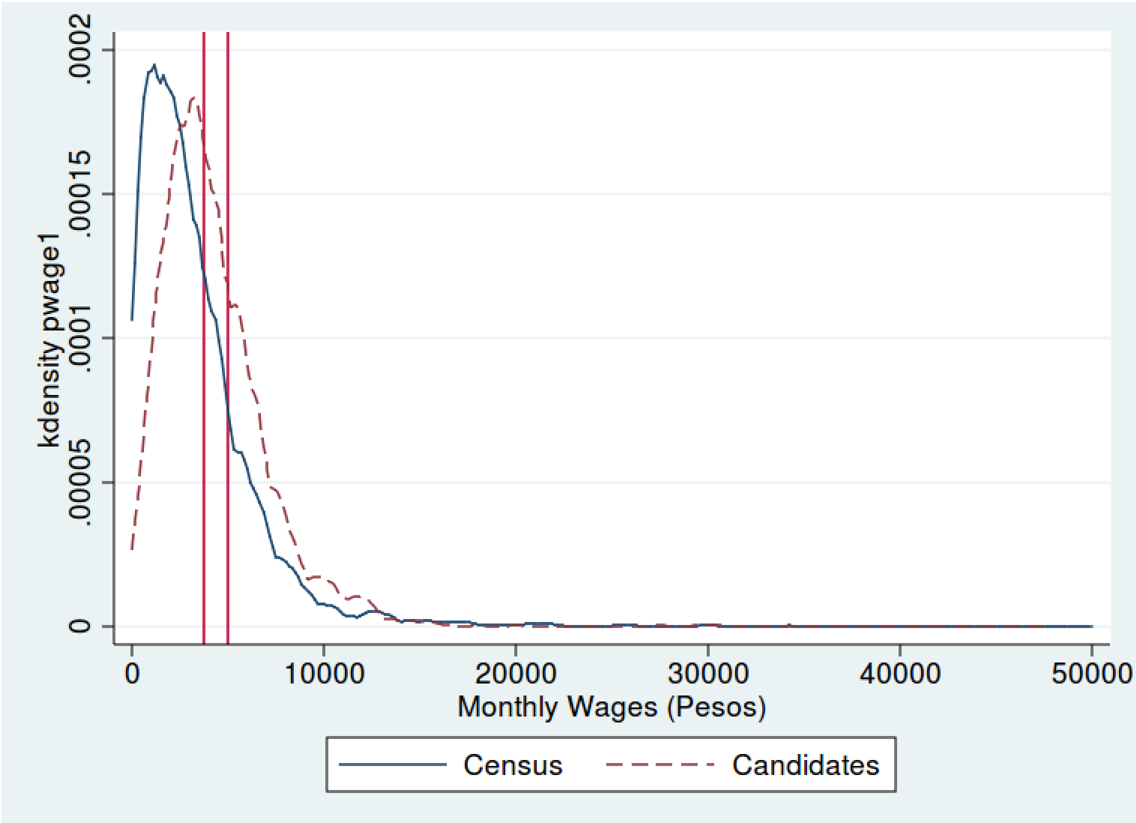

FIGURE A2: DISTRIBUTIONS OF WAGES

Figure Notes: This figure plots the distributions of wages for the applicant pool (dashed line) and the population (solid line) in the 10 regions in which the program is operating. Each density was estimated using an Epanechnikov kernel and an optimal bandwidth. The vertical lines denote the experimental wage offers of 3,750 and 5,000 pesos. The wage data for the population come from the 2010 population census.

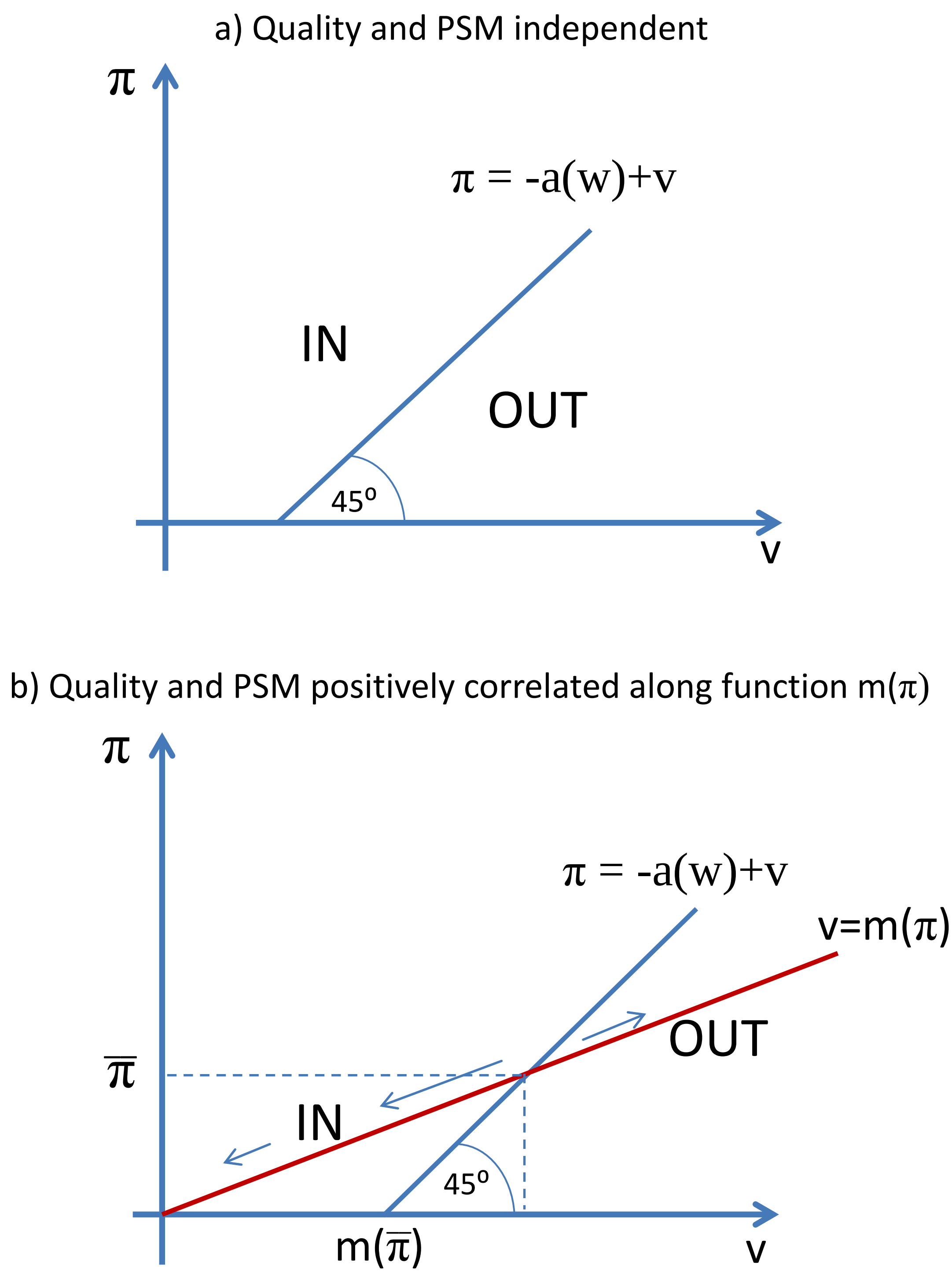

FIGURE A3: SELECTION PATTERNS IN THE QUALITY-PSM SPACE

Table A1. Comparison of RDP Municipalities to the Rest of Mexico

	Non RDP municipalities (1)	RDP municipalities (2)	Difference (3)	Standard error (4)
Population	43659	20456	23202	3306
Infant mortality (deaths x 1000 births)	22.36	32.73	10.37	0.84
Literacy rate (% of literate 15 year olds and older)	84.19	69.64	14.54	0.91
Human Development Index 2005	0.76	0.67	0.09	0.00
Income per capita 2005 (monthly, Pesos)	6148.55	3663.75	2484.80	120.26
Drug cartel is present	0.29	0.37	-0.08	0.04
Subversive group is present	0.04	0.50	-0.45	0.04
Drug-related deaths per 1000 inhabitants	19.12	16.32	2.80	4.58
Altitude variation (standard deviation)	192.28	340.46	-148.18	12.46
Average annual precipitation (mms)	1060.14	1278.91	-218.77	33.17
Observations	2289	167		

Notes: This table compares the mean socio-economic characteristics of the municipalities in the program to those not in the program. Column (1) reports the mean of the corresponding variables among municipalities that are not in the program. Column (2) reports the mean of the corresponding variables among municipalities in the program. Column (3) reports the difference in the mean and column (4) reports the standard error of the difference. Demographic and socioeconomic data (Population, Literacy, Infant mortality, Human Development Index, Income per capita) are from the 2005 population census; Confidential government data includes: Indicator variables for Drug cartel is present - equals 1 if the municipality has at least one drug cartel operating in it, zero otherwise; Subversive group is present - equals 1 if the municipality has at least one subversive group operating in it; Drug-related deaths per 1000 inhabitants. Data obtained from the Instituto de Estadística y Geografía (INEGI) include: Altitude variation - standard deviation of altitudes in the municipality, in meters and Average annual precipitation (in milimeters).

Table A2. The Correlates of Previous Earnings

Dependent variable	Log wages							Wages
Data source	Applicants' data				MxFLS	Applicants' data		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	0.193	0.141	0.174	0.181	0.132	0.178	0.188	585.822
	[0.040]***	[0.049]***	[0.050]***	[0.050]***	[0.066]**	[0.041]***	[0.041]***	[135.633]***
Experience	0.047	0.047	0.048	0.050	0.077	0.048	0.045	296.641
	[0.008]***	[0.008]***	[0.008]***	[0.008]***	[0.010]***	[0.008]***	[0.008]***	[29.821]***
Experience^2	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-5.419
	[0.000]***	[0.000]***	[0.000]***	[0.000]***	[0.000]***	[0.000]***	[0.000]***	[1.139]***
Education	0.078	0.078	0.078	0.073	0.082	0.071	0.072	287.282
	[0.009]***	[0.009]***	[0.009]***	[0.009]***	[0.011]***	[0.010]***	[0.010]***	[32.790]***
Height		0.390	0.168	0.071	0.561			
		[0.209]*	[0.219]	[0.221]	[0.311]*			
Indigenous			-0.134	-0.120	-0.119	-0.115	-0.106	-645.339
			[0.041]***	[0.041]***	[0.114]	[0.040]***	[0.041]**	[132.242]***
IQ				0.025	0.055	0.022	0.022	93.558
				[0.008]***	[0.011]***	[0.008]***	[0.008]***	[26.201]***
Big 5 Personality Traits								
Extrovert						-0.092	-0.078	-122.038
						[0.039]**	[0.038]**	[125.622]
Agreeable						-0.029	-0.008	-73.821
						[0.053]	[0.053]	[167.851]
Conscientious						0.011	0.003	-111.196
						[0.054]	[0.053]	[224.794]
Neurotic						-0.119	-0.111	-361.174
						[0.048]**	[0.046]**	[187.351]*
Open						0.026	0.032	413.209
						[0.044]	[0.044]	[149.963]***
Number of observations	1433	1433	1433	1433	1569	1433	1433	2006
R-squared	0.11	0.11	0.12	0.13	0.13	0.13	0.16	0.18
Joint significance of Big 5 (p-value)	n/a	n/a	n/a	n/a	n/a	0.08	0.12	0.04
Region intercepts	N	N	N	N	N	N	Y	N

Notes: This table reports estimates from OLS regressions. In columns (1)-(4) and (6)-(7), the dependent variable is the monthly wage the candidates reported in their previous jobs, expressed in logarithms. In column (8), the dependent variable is the candidate's wage in the previous job in levels, where missing wages have been replaced with zero. In column (5) the dependent variable is the current monthly wage of respondents to the MxFLS 2005 survey. The regression presented in column (5) has been reweighted so that the observable characteristics in the MxFLS sample match those in the applicant data. The experience variable defined by subtracting 6 and the years of schooling from the person's age. See the data appendix for more information on the variables including their sources. * indicates statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Robust standard errors are reported in brackets.

Table A3: Fill-in Errors by Treatment Status

Question topic	Control	Treatment Effect
	(1)	(2)
Date of birth	0.016	-0.010
		[0.006]
Gender and Civil Status	0.012	-0.002
		[0.005]
Education level	0.151	0.014
		[0.028]
Parent's education	0.061	-0.018
		[0.012]
Household characteristics	0.427	-0.016
		[0.032]
Religion	0.052	0.011
		[0.012]
Identification	0.170	0.025
		[0.058]
Total	0.948	-0.025
		[0.089]

Notes: This table estimates the effects of higher wages on the number of mistakes made by the applicant in filling in the questionnaire for a subset of questions. Each row is a separate regression using the variable listed as the dependent variable. Column (1) reports the mean of the variable in the control group (low wage announcement), column (2) reports the coefficient on the treatment variable in a regression that includes region intercepts. The variable `Household characteristics' refers to questions about family size, number of kids, number of family members living abroad, and head of household status. `Identification' refers to 6 questions about possession of different identification cards (e.g. passport, driver's license, etc.) * indicates statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Clustered standard errors at the level of the locality are reported in brackets.

Table A4. The Correlates of Pro-social Behavior

Dependent variable	Charity (1)	Volunteer (2)	Belongs to a political party (3)	Voted (4)	Altruism (5)	Negative reciprocity (6)	Cooperation (7)	Importance of wealth (8)
PSM index	0.076	0.070	0.020	-0.010	1.242	-0.065	1.311	-0.154
	[0.019]***	[0.017]***	[0.011]*	[0.015]	[0.282]***	[0.019]***	[0.405]***	[0.054]***
Raven score	-0.008	0.004	-0.009	-0.002	-0.280	-0.010	-0.065	0.026
	[0.005]*	[0.004]	[0.003]***	[0.004]	[0.076]***	[0.005]**	[0.104]	[0.014]*
Raven score × (PSM index-E(PSM index))	0.004	0.000	0.002	0.012	-0.306	-0.006	0.260	-0.022
	[0.006]	[0.006]	[0.004]	[0.005]**	[0.118]***	[0.006]	[0.142]*	[0.019]
Big 5 index	0.042	0.046	0.001	0.052	-0.112	0.010	-0.316	-0.056
	[0.019]**	[0.017]***	[0.011]	[0.015]***	[0.255]	[0.019]	[0.401]	[0.053]
Male	-0.001	0.062	0.023	-0.015	-0.639	0.001	1.751	0.210
	[0.023]	[0.021]***	[0.013]*	[0.018]	[0.326]**	[0.023]	[0.490]***	[0.067]***
Years of schooling	-0.001	0.006	-0.002	0.044	-0.002	-0.006	0.080	0.015
	[0.005]	[0.005]	[0.003]	[0.004]***	[0.073]	[0.005]	[0.107]	[0.014]
Age	-0.001	0.003	0.004	0.018	-0.013	0.000	0.050	-0.003
	[0.002]	[0.001]**	[0.001]***	[0.001]***	[0.028]	[0.002]	[0.040]	[0.005]
Number of observations	1941	1942	1945	1944	1932	1934	1907	1819
R-squared	0.02	0.04	0.02	0.16	0.03	0.01	0.02	0.02

Notes: This table reports estimates from OLS regressions, where the dependent variable is as indicated in each column. See the data appendix for more information on the variables including their sources. * indicates statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Robust standard errors are reported in brackets.

Table A5: The Effects of Financial Incentives on Recruitment by Applicant Characteristics

Dependent variable	Acceptance					
	(1)	(2)	(3)	(4)	(5)	(6)
High wage offer	0.160	0.204	0.252	0.150	0.156	0.299
	[0.080]**	[0.066]***	[0.093]***	[0.053]***	[0.054]***	[0.125]**
High wage offer × High IQ	-0.016					-0.008
	[0.112]					[0.113]
High IQ	-0.010					-0.023
	[0.091]					[0.093]
High wage offer × High outside wage		-0.251				-0.237
		[0.153]				[0.164]
High outside wage		0.169				0.157
		[0.138]				[0.141]
High wage offer × Male			-0.177			-0.164
			[0.121]			[0.119]
Male			0.027			0.035
			[0.097]			[0.090]
High wage offer × (Big 5 index-E(Big 5 index))				-0.052		0.064
				[0.070]		[0.091]
Big 5 index				0.053		-0.060
				[0.056]		[0.075]
High wage offer × (PSM index-E(PSM index))					0.015	-0.058
					[0.084]	[0.082]
PSM Index					-0.019	0.076
					[0.068]	[0.067]
Observations	350	350	350	349	344	344
R-squared	0.1	0.11	0.12	0.13	0.13	0.13

Notes: This table estimates whether the effects of higher wages on conversion (a selected candidate filling a vacancy) vary with applicant characteristics. In all columns, the dependent variable is an indicator equal to 1 if the person accepted the offer, zero if the applicant rejected the offer or could not be reached. In addition to the controls listed in the table, all regressions included intercepts indicating a region by indigenous requirement pair. See the data appendix for more information on the variables including their sources. * indicates statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. Clustered standard errors at the level of the locality are reported in brackets.