On-Line Appendix C for

Inequality, the Urban-Rural Gap and Migration*

Alwyn Young
Department of Economics
London School of Economics
This Draft: June 2013

*Appendices A and B in the published paper.

Appendix C: Extending the Model

In this appendix I show how the model can be modified to make the migration decision determinate and account for differences in the urban residence probabilities of the urban and rural born. The later, in particular, produces predictions regarding the consumption differences of the urban and rural born living in the same region.

(a) Determinate Migration Decisions

For the quantity of skilled and unskilled labour input used in sector i, as in the production function (6) in the text, substitute:

$$\text{(C1)}\quad S_i = \int_{u \in SetS_i} z_S^i(u)\,,\ US_i = \int_{u \in SetUS_i} z_{US}^i(u)$$

where $SetS_i$ and $SetUS_i$ are the sets of skilled and unskilled workers who choose to work in industry i and $z_x^i(u)$ represents the efficacy of person u with skills x when working in sector i.[1] When an individual's education is completed and their skill status determined by $P(E)$, as in the text, they are endowed with a paired set of sectoral productivities (z_S^U, z_S^R) or (z_{US}^U, z_{US}^R), depending upon whether they are skilled or unskilled. These paired productivities are independent draws from the cumulative distribution functions $G_x^i(z)$, $i = U, R$ and $x = S, US$, and determine where an individual chooses to live and work. With w_S^i and w_{US}^i denoting the offered skilled and unskilled wage per unit of effective labour (z) in sector i, an individual with skills x chooses sector i over j if $w_x^i z_x^i > w_x^j z_x^j$. Thus, for example, the probability a skilled individual chooses to work and live in the urban sector (Π_S^U) is given by

$$\text{(C2)}\quad \Pi_S^U = \int_{\underline{z}_S^U}^{\bar{z}_S^U} G_S^R(w_s^U v / w_s^R)\, dG_S^U(v)$$

where [$\underline{z}_S^U, \bar{z}_S^U$] is the support of z_S^U.

I will now prove a few characteristics of the equilibrium. First, I will show that if the distribution functions determining the efficacy draws follow the common form $G_x^i(z) = G(z/\lambda_x^i)$, then $\alpha_U > \alpha_R$ implies $\Pi_S^U > \Pi_{US}^U$. Thus, provided there is enough similarity in the distributions of the efficacy draws, the higher skill intensity of the urban sector guarantees that skilled workers are more likely than unskilled workers to reside there. As the reader will recall, these residence probabilities were what determined the urban-rural gap in the text. Second, I will show that $P'(E) > 0$ and $\Pi_S^U > \Pi_{US}^U$ are enough to guarantee that rural to urban migrants are drawn from the upper end of the educational distribution of the rural born and urban to rural migrants are drawn from the lower end of the educational distribution of the urban born, as was shown in

[1]Where I have to denote both the sector and the type of worker in a single term, I use the subscript to denote the type of worker and the superscript the sector.

Table II earlier. Finally, I note that if the reader goes back to the presentation of the urban-rural gap in the text, she will see that it was predicated on the notion that there was a common ln skilled and ln unskilled wage across industries, so to compare earnings across sectors it was sufficient to compare the probability a worker was skilled. The equivalent measure in this version of the model is average ln earnings, equal to the expectation of $\ln(w_x^i z_x^i)$ conditional on a worker residing in a sector. While for the marginal worker potential earnings in the two sectors are identical, average earnings or ln earnings by type of worker depend upon the inframarginal distribution of efficacy. For efficacy draws from Fréchet distributions (as proven further below), average earnings and ln earnings by worker type are equalized across sectors, but this is not generally true for any arbitrary distribution of efficacy. Thus, in this version of the model one must assume that differences across sectors in average ln earnings by worker type are small relative to average differences between skilled and unskilled workers. This issue aside, the presentation of the urban-rural gap is exactly the same for this model as it was for the simpler framework in the text.

<u>Proposition I</u>: Let $G_x^i(z) = G(z/\lambda_x^i)$. If $\alpha_U > \alpha_R$, then $\Pi_S^U > \Pi_{US}^U$ is assured.

Proof: I begin by noting that

$$\text{(C3)}\quad \Pi_x^i = \int_{\lambda_x^i z_l}^{\lambda_x^i z_u} G_x^j(w_x^i v / w_x^j)\, g_x^i(v)\, dv \;=\; \int_{z_l}^{z_u} G(\tilde{w}_x^i \tau)\, g(\tau)\, d\tau \;=\; h(\tilde{w}_x^i),$$

where z_l and z_u represent the lower and upper bounds of the support of $z \sim G(z)$ (including, possibly, 0 and ∞), lower case g's denote densities with $g_x^i(v) = g(v/\lambda_x^i)/\lambda_x^i$, and where I have made the substitution $\tau = v/\lambda_x^i$, and $\tilde{w}_x^i = w_x^i \lambda_x^i / w_x^j \lambda_x^j$. Clearly, *h(.)* is monotonically increasing in its argument. The total labour input of type *x* is given by

$$\text{(C4)}\quad Z_x^i = \int_{\lambda_x^i z_l}^{\lambda_x^i z_u} v\, G_x^j(w_x^i v / w_x^j)\, g_x^i(v)\, L_x\, dv,$$

where L_x is the total number of individuals of type *x*, so $w_S^i Z_S^i / w_{US}^i Z_{US}^i = \alpha_i/(1-\alpha_i)$ implies

$$\text{(C5)}\quad k(\tilde{w}_S^U) = \frac{\alpha_U/(1-\alpha_U)}{\alpha_R/(1-\alpha_R)} k(\tilde{w}_{US}^U),$$

where

$$\text{(C6)}\quad k(\tilde{w}_x^i) = \frac{w_x^i Z_x^i}{w_x^j Z_x^j} = \frac{w_x^i \int_{\lambda_x^i z_l}^{\lambda_x^i z_u} v\, G_x^j(w_x^i v / w_x^j)\, g_x^i(v)\, L_x\, dv}{w_x^j \int_{\lambda_x^j z_l}^{\lambda_x^j z_u} v\, G_x^i(w_x^j v / w_x^i)\, g_x^j(v)\, L_x\, dv} = \frac{\tilde{w}_x^i \int_{z_l}^{z_u} \tau\, G(\tilde{w}_x^i \tau)\, g(\tau)\, d\tau}{\int_{z_l}^{z_u} \tau\, G(\tau/\tilde{w}_x^i)\, g(\tau)\, d\tau}$$

and where I have used the substitutions $\tau = v/\lambda_x^i$ and $\tau = v/\lambda_x^j$ in the numerator and denominator. *k(.)* is monotonically increasing in its argument. Consequently, from (C5) we see that $\tilde{w}_S^U > \tilde{w}_{US}^U$, which through (C3) implies $h(\tilde{w}_S^U) > h(\tilde{w}_{US}^U)$ or $\Pi_S^U > \Pi_{US}^U$, thereby completing the proof. In general, the assumption that $\alpha_U > \alpha_R$ will work to ensure that skilled workers are more likely to choose to work in urban areas than unskilled workers, but this cannot be guaranteed if one allows the distribution functions of the productivity draws z_x^i to take radically different forms.

Proposition II: Let $\overline{E}_{ij}$ equal the mean educational attainment of individuals born in region *i* working in region *j*. With *P'(E) > 0* and $\Pi_S^U > \Pi_{US}^U$, $\overline{E}_{iU} > \overline{E}_{iR}$ is assured.

Proof: I begin by defining:

$$\text{(C7)}\quad \overline{E}_{ij} = \int_{E_l}^{E_u} E\, dG_i^j(E) \quad \text{where} \quad G_i^j(E) = \frac{\int_{E_l}^{E} \Pi_E^j \, dG_i(E)}{\int_{E_l}^{E_u} \Pi_E^j \, dG_i(E)}$$

$$\text{and} \quad \Pi_E^j = P(E)\Pi_S^j + (1-P(E))\Pi_{US}^j$$

and where E_l and E_u are the lower and upper bounds of the support of *E*, $G_i^j(E)$ is the cumulative distribution function of the probability an individual born in region *i* residing in region *j* has educational attainment *E*, $G_i(E)$ is the cumulative distribution function of the educational attainment of individuals born in region *i* and Π_E^j is the probability an individual with educational attainment *E* resides in region *j*. To prove the proposition, it is enough to show that $G_i^U(E) < G_i^R(E)\ \forall\ E$. Using $\Pi_E^R = 1 - \Pi_E^U$, this amounts to showing that:

$$\text{(C8)}\quad \frac{\int_{E_l}^{E} \Pi_E^U \, dG_i(E)}{\int_{E_l}^{E_u} \Pi_E^U \, dG_i(E)} < \frac{\int_{E_l}^{E} [1-\Pi_E^U] dG_i(E)}{\int_{E_l}^{E_u} [1-\Pi_E^U] dG_i(E)} = \frac{G_i(E) - \int_{E_l}^{E} \Pi_E^U \, dG_i(E)}{1 - \int_{E_l}^{E_u} \Pi_E^U \, dG_i(E)}.$$

Cross-multiplying and canceling terms, we need to show that

$$\text{(C9)}\quad \frac{\int_{E_l}^{E} \Pi_E^U \, dG_i(E)}{G_i(E)} < \int_{E_l}^{E_u} \Pi_E^U \, dG_i(E).$$

The right-hand side is the mean value of Π_E^U, the left-hand side is the truncated mean value of Π_E^U. With *P'(E) > 0* & $\Pi_S^U > \Pi_{US}^U$, we know that Π_E^U is increasing in *E*. Consequently, its right truncated mean is less than its mean. This establishes that $G_i^U(E) < G_i^R(E)\ \forall\ E$, and hence $\overline{E}_{iU} > \overline{E}_{iR}$, the mean educational attainment of an individual born in region i residing in urban areas is always greater than that of individuals from the same region residing in rural areas. Because of this, rural to urban migrants are better educated than rural permanent residents and urban to rural migrants are worse educated than urban permanent residents.

Proposition III: Let $G_x^i(z) = G(z/\lambda_x^i)$, where *G(z)* is of the Fréchet form $exp(-z^{-\theta})$, with $\theta > 0$. Then $E(f(w_x^i z_x^i(u)) \mid u \in Setx_i) = E(f(w_x^j z_x^j(u)) \mid u \in Setx_j)$.

Proof: I simply note that

$$\text{(C10)}\quad E(f(w_x^i z_x^i(u))\,|\,u \in Setx_i) = \frac{\int_0^\infty f(w_x^i v)\, G_x^j(w_x^i v / w_x^j)\, g_x^i(v)\, L_x\, dv}{\int_0^\infty G_x^j(w_x^i v / w_x^j)\, g_x^i(v)\, L_x\, dv}$$

$$= \frac{\int_0^\infty f(w_x^i v)\exp(-(w_x^i v/\lambda_x^j w_x^j)^{-\theta})\exp(-(v/\lambda_x^i)^{-\theta})\theta \lambda_x^{i\,\theta} v^{-\theta-1} L_x\, dv}{\int_0^\infty \exp(-(w_x^i v/\lambda_x^j w_x^j)^{-\theta})\exp(-(v/\lambda_x^i)^{-\theta})\theta \lambda_x^{i\,\theta} v^{-\theta-1} L_x\, dv}$$

$$= \frac{\int_0^\infty f(\tau)\exp(-(\tau/\lambda_x^j w_x^j)^{-\theta})\exp(-(\tau/\lambda_x^i w_x^i)^{-\theta})\tau^{-\theta-1}\, d\tau}{\int_0^\infty \exp(-(\tau/\lambda_x^j w_x^j)^{-\theta})\exp(-(\tau/\lambda_x^i w_x^i)^{-\theta})\tau^{-\theta-1}\, d\tau}$$

where I have substituted using $\tau = w_x^i v$ and cancelled $\theta L_x (w_x^i \lambda_x^i)^\theta$ from the numerator and the denominator. As the $\lambda_x^i w_x^i$ and $\lambda_x^j w_x^j$ terms in the integrals are symmetric, it follows that $E(f(w_x^i z_x^i(u))\,|\,u \in Setx_i) = E(f(w_x^j z_x^j(u))\,|\,u \in Setx_j)$, completing the proof.

Letting $f(w_x^i z_x^i) = w_x^i z_x^i$ or $\ln(w_x^i z_x^i)$, we see that expected wages and ln wages by worker type equalize across sectors, as claimed earlier above. Moreover, using $w_x^i Z_x^i = E(w_x^i z_x^i)\Pi_x^i L_x$ and $w_S^i Z_S^i / w_{US}^i Z_{US}^i = \alpha_i/(1-\alpha_i)$, we have:

$$\text{(C11)}\quad \frac{\alpha_U/(1-\alpha_U)}{\alpha_R/(1-\alpha_R)} = \frac{E(w_S^U z_S^U)\Pi_S^U L_S \big/ E(w_{US}^U z_{US}^U)\Pi_{US}^U L_{US}}{E(w_S^R z_S^R)\Pi_S^R L_S \big/ E(w_{US}^R z_{US}^R)\Pi_{US}^R L_{US}} = \frac{\Pi_S^U/(1-\Pi_S^U)}{\Pi_{US}^U/(1-\Pi_{US}^U)}$$

This is identical to equation (8) in the paper. To the degree that average earnings by worker type equalize across sectors, the odds ratio of the urban residence probabilities are completely determined by the "odds ratio" of the factor shares. The absolute residence probabilities, however, are always determined by other aspects of the equilibrium, like the demand for urban and rural products and the educational attainment of the population.

(b) Differences in Residence Probabilities & Intra-Regional Consumption

Table C1 reports the educational characteristics and population shares of different migrant groups and compares them to the values implied by a simple residence equation estimated off of all households using educational attainment and a constant alone, i.e. the equation used to predict the urban residence probability of people with low and high educational attainment for the regressions in Table VIII in the text. As shown, this equation does an excellent job of predicting the educational characteristics of different groups. In other words, if one applies the nationwide average urban residence probability by educational attainment to the urban and rural born, taking

Table C1: Population Characteristics: Data and Model (34 countries)								
	Mean Educational Attainment				Population Shares			
	E_{RU}	E_{RR}	E_{UU}	E_{UR}	S_{RU}	S_{RR}	S_{UU}	S_{UR}
Data Mean	5.33	3.32	7.32	5.04	.226	.774	.780	.220
Model Mean	4.97	3.06	8.42	5.10	.365	.635	.521	.479
Corr: Model & Data	.994	.996	.959	.981	.928	.928	.744	.744

Notes: The two subscripts denote an origin-destination combination. Thus, E_{RU} and S_{RU} are the mean educational attainment and population share of rural born households residing in urban areas. Calculations are for the 34 country averages across 64 surveys that allow origin-destination breakdown of households based upon data on women's residence prior to the age of 12, as described earlier in Section II. Patterns for origin-destination breakdown based upon men's data are similar.

as given their educational attainment, one gets a good approximation of the educational characteristics of those who migrate and those who stay at home. However, as the right hand panel of the table shows, this simple equation over-predicts the share of the rural population moving to urban areas and the share of the urban population moving to rural areas. This indicates that there is a tendency for individuals of a given educational attainment to stay in their region of birth. Three ways in which the model can be modified to allow for this characteristic come to mind:[2] (1) by introducing a real cost (not barrier) to moving in either direction; (2) by specifying that, conditional on their skilled or unskilled status, individuals are more likely to acquire abilities used by industries in their region of birth;[3] and (3) by specifying that the quality of education, i.e. the probability of acquiring skill for a given level of educational attainment, varies by region of birth. This last option produces predictions that are quite similar in form to those discussed in the text, so I explore it here.

One can use a discrete choice equation to estimate the probability a household lives in urban areas as a function of their educational attainment *and* their region of birth. This produces a predictive index equal to $\beta_E E + \Delta$, where β_E is the coefficient on educational attainment and Δ is a dummy coefficient for urban birth. This is justified, within the context of the model, by arguing that an individual receiving E years of education in urban areas has the same probability of being

[2]Strictly speaking, since migration is indeterminate in the model of the text, it has no specific predictions regarding the characteristics and number of migrants. However, if one endogenizes the migration decision using the framework described above, assuming that individuals raised in urban and rural areas share a common P(E) and common distribution functions for the sectoral efficacy draws z, then the model indicates that a national residency equation should be used to predict the characteristics and number of migrants, i.e .conditional on educational attainment the residency decision is independent of region of birth.

[3]In the framework outlined above one could specify a rightward shift of the distribution for individuals born in the region (i.e. λ_x^i in $G_x^i(z) \sim G(z/\lambda_x^i)$ is greater if the individual is born in region i).

skilled as someone receiving $E + \Delta$ years of education in rural areas. Thus, the probability urban or rural born individuals live in urban areas is given by

$$\text{(C12)}\quad \Pi_E^U(\text{Uborn}) = P(E+\Delta)\Pi_S^U + (1-P(E+\Delta))\Pi_{US}^U$$

$$\Pi_E^U(\text{Rborn}) = P(E)\Pi_S^U + (1-P(E))\Pi_{US}^U$$

I note, in passing, that with a dummy for region of birth the predicted population shares by migrant status (as in Table C1) automatically match the data.

The difference between the consumption of urban and rural born individuals residing in the same region *i* (the "within gap") is given by:

$$\text{(C13)}\quad \text{WithinGap} = \ln(w_s)\left[\frac{\Pi_S^i P(\mu_E^i+\Delta)}{\Pi_S^i P(\mu_E^i+\Delta)+[1-P(\mu_E^i+\Delta)]\Pi_{US}^i} - \frac{\Pi_S^i P(\mu_E^i)}{\Pi_S^i P(\mu_E^i)+[1-P(\mu_E^i)]\Pi_{US}^i}\right]$$

where, using the notation and framework of the text (not that of the section above) w_S is the economy-wide wage for undifferentiated skilled workers and where I calculate the measure at the mean level of educational attainment in the region *i*, μ_E^i, as these estimates are by and large determined by the mean regional household. Using the fact that R_E still equals $ln(w_S)P'(E)$, and linearizing around $\Delta = 0$ and (as in the text) $\Pi_S^i = \Pi_{US}^i = \overline{\Pi}$ and $\mu_E^i = \overline{\mu}$, we have:

$$\text{(C14)}\quad \text{WithinGap} = 1*\Delta + 0*(\Pi_S^i - \overline{\Pi}) + 0*(\Pi_{US}^i - \overline{\Pi}) + 0*(\mu_E^i - \overline{\mu})$$

In regressions, I will take the residence probabilities of rural born individuals with zero years of education (those with the lowest measureable human capital) as proxies for the residence probability of the unskilled, the residence probabilities of urban born individuals with 16 years of education (the opposite extreme) as proxies for the residence probability of the skilled, and the empirically estimated mean regional educational attainment and urban-born dummy in the residence equation as measures of μ_E^i and Δ. Once again the model has strong predictions. Since the within gap is identically equal to 0 when $\Delta = 0$, none of the other regressors matters and the constant term in the regression is zero, while the coefficient on Δ should equal 1. I note in passing that (C14) implies that the differences between urban and rural born households living in urban areas and the differences between urban and rural born households living in rural areas should be equal. This restriction is not rejected at the 1% level for 19 of the 33 country estimates based on women's migration status and 23 of the 25 country estimates based on men's migration status in Table VII in the text.

Before turning to the empirical results, I note that for the model just described the urban-rural gap is given by

$$\text{(C15)}\quad \text{URGap} = \ln(w_s)\left[\frac{\Pi_S^U A}{\Pi_S^U A + \Pi_{US}^U B} - \frac{\Pi_S^R A}{\Pi_S^R A + \Pi_{US}^R B}\right]$$

$$\text{where}\quad A = L_E^U P(\mu_E + \Delta) + L_E^R P(\mu_E) \quad \text{and} \quad B = L_E^U[1 - P(\mu_E + \Delta)] + L_E^R[1 - P(\mu_E)]$$

and where L_E^i is the number of workers of educational attainment *E* born in region *i*. Linearizing, we have:

$$\text{(C16)}\quad \text{UR Gap} \;=\; \beta*\left(\Pi_S^U - \overline{\Pi}\right) - \beta*\left(\Pi_{US}^U - \overline{\Pi}\right) + \;0*\Delta \;+\; 0*(\mu_E - \overline{\mu})$$

$$\text{where}\quad \beta \;=\; \frac{P(\overline{\mu})[1 - P(\overline{\mu})]}{P'(\overline{\mu})\overline{\Pi}(1 - \overline{\Pi})}$$

Since the urban-rural gap is still identically zero when $\Pi_S^U = \Pi_{US}^U$, the results of the text follow, with the additional implication that the urban-born residence dummy has no influence.

In Table C2 I report results with household urban residence probabilities and within region consumption gaps estimated, separately, on the basis of the migrant status of women and men in the household.[4] As shown in columns (1) and (4), the addition of the dummy for urban born (Δ) in the residence equation does not change the conclusions of the text. The null hypothesis in the urban-rural gap regression that the coefficients on the skilled and unskilled urban residence probability are opposite in sign and equal in magnitude, and that all of the remaining coefficients, including the constant term, are zero has a very large p-value, i.e. is nowhere near being rejected. The samples in these regressions are much smaller than those used in Table VIII of the text because, with the inclusion of Δ, the analysis here is restricted to countries with migration data. The remaining columns of the table report regressions for within region consumption differences. In this case, the model's prediction is that the coefficient on the residence dummy for urban born is one and all of the remaining terms, including the constant, are zero. The results here are mixed. The null hypothesis is not rejected for the rural within gap, but is poorly received in the regressions for the urban within gap. Thus, in its predictions regarding the equality of urban and rural within region consumption differences (see above) and the determinants of the urban-rural and within region consumption gaps, this approach to explaining differences in the urban residence probability by region of birth gets some traction, but is by no means an unqualified success.

[4]As in the text, since all of the consumption data are at the household level, I use the household as the unit of analysis classifying households into native born or migrants based upon where all of the male or female members lived prior to the age of 12. I use the male-based residence equations for the male-based consumption equations and the female-based residence equations for the female-based consumption equations (see Tables VII & VIII earlier).

Table C2: Residual Consumption Gaps as Functions of Residence Probabilities						
	(1)	(2)	(3)	(4)	(5)	(6)
	UR Gap	Rural Within Gap	Urban Within Gap	UR Gap	Rural Within Gap	Urban Within Gap
	Based on women's migration status (33 country observations)			Based on men's migration status (25 country observations)		
α: constant	-13.9 (8.02)	3.43 (3.54)	-.113 (2.23)	-7.93 (8.30)	2.90 (3.53)	-2.21 (2.72)
β_1: $\Pi_S^U - \overline{\Pi}$	35.6 (13.4)	-4.65 (6.05)	-.212 (3.76)	19.1 (13.4)	-2.35 (5.72)	2.36 (4.41)
β_2: $\Pi_{US}^U - \overline{\Pi}$	-18.8 (7.53)	-.573 (3.25)	-1.22 (2.04)	-22.1 (13.1)	.113 (5.52)	-7.65 (4.29)
β_3: $\mu_E - \overline{\mu}$	.154 (.223)	-.081 (.106)	.126 (.073)	.182 (.336)	-.065 (.156)	-.092 (.130)
β_4: Δ	-.733 (.789)	.559 (.349)	.502 (.221)	-.027 (1.37)	.264 (.563)	.007 (.456)
p-value on H_0	.480	.660	.000	.880	.660	.022

Notes: Π_S^U, Π_{US}^U = urban residence probability of the skilled and unskilled, proxied by the urban residence probabilities of urban individuals with16 years of education and rural individuals with 0 years of education, respectively. μ_E = estimated mean household educational attainment in the country (UR Gap) or the region (Within Gap). $\overline{\Pi}$ and $\overline{\mu}$ = average urbanization rate and mean household educational attainment of the sample. Δ = dummy for urban birth in logit model of urban residence as function of educational attainment and region of birth. H_0 URGap: $\beta_1 = -\beta_2$ and $\alpha = \beta_3 = \beta_4 = 0$. H_0 Within Gap: $\alpha = \beta_1 = \beta_2 = \beta_3 = 0$ and $\beta_4 = 1$. Country sample is those meeting migration sample criteria discussed in text surrounding Table II.