Moving from a Poor Economy to a Rich One: A Natural Experiment

Eran Yashiv
Tel Aviv University, CReAM (UCL), and CEPR

June 2, 2019*

Abstract

The phenomenon of workers moving from poor to rich economies is a very prevalent one. Understanding the gains inherent in such a move faces the difficulty of disentangling the pure effects of income differences from many other determinants. The paper addresses this difficulty using a unique natural experiment: the case of Palestinian workers from the West Bank and from Gaza working in Israel. During most of the 1980s a sizeable fraction of the male labor force from these areas worked in Israel, a far richer economy. The set up was such that determinants for moving other than income differences played almost no role. A worker could decide to work in this richer economy and do so by a weekly commute. Hence this paper is able to estimate the moving decision without omitted variables bias.

The empirical work caters for key facts concerning rich and poor countries differences characterized by the recent development accounting literature. The key findings are that productivity differences in favor of the richer economy, due to differences in TFP and in capital, operate to raise the wages of movers. Lower returns to human capital for movers, working in low-skill occupations, operate to offset these effects.

Key Words: movers, stayers, rich and poor economies, natural experiment, development accounting, TFP differentials, human capital differences, selection.

JEL Codes: E24, J24, O15.

* *I am grateful to David Autor, Jim Heckman, and John Kennan for useful conversations, and to seminar participants in numerous places for useful feedback and suggestions on previous versions. Any errors are my own. The graphs in this paper are best viewed in color.*

1 Introduction

The phenomenon of workers moving from a poor to a rich economy is a very prevalent one. It may be an internal migration or commuting move[1] or migration across countries. When a worker moves to an economy richer than the home economy, what is gained by the move? It is not easy to answer this question, given the difficulty to disentangle the effects of income differences from many other determinants of such mobility. The set of determinants includes geographical distance, socio-demographic factors including family linkages and social networks, credit constraints, welfare benefits, insurance motives, psychological issues,[2]and more. Moreover, one needs to address the question of what workers newly experience in the richer economy (say, higher productivity), what is taken from the poorer economy (e.g., human capital), and choices (self-selection).

This paper studies a unique natural experiment that allows to isolate the "pure" effect of income differences and to cater for these issues. This is the case of Palestinian workers from the West Bank and from Gaza working in Israel. During most of the 1980s a sizeable fraction of the male labor force from these areas worked in Israel, a far richer economy. The features of this labor market, delineated below, were such that the other cited factors played almost no role. There thus existed a special situation, whereby a worker could decide on work in a richer economy and place himself there by a weekly commute. Without the confounding factors, the decision to work in the richer economy can thus be estimated without bias. This paper is able to answer the following questions: what happens when a worker moves to a richer economy, without altering 'embodied' skills, such as those gained through education or experience? What are the self-selection mechanisms in operation, when the motivation is purely income differences?

I use a model catering for notable facts concerning rich and poor countries income differences characterized by the recent development accounting literature. The latter suggests sizeable rich-poor countries income differences, while debating the relative weights of various constituents. One consequence for moving decisions, highlighted by recent papers, is that there could potentially be very large gains due to the afore-cited large income differences.

Specifically, the paper addresses four questions, pertaining to the wage differentials motivating movers from poor to rich economies. First, what is the role of the productivity advantage of the richer, host economy relative

[1]Thus, for example, using data from 170 Demographic and Health Surveys for 65 countries, Young (2013) finds that about one out of every four or five individuals raised in rural areas migrates to urban areas as a young adult.

[2]For example, Kennan and Walker (2011) show that attachment to home is an important determinant of internal migration decisions in the U.S.

to the poorer, home economy? Second, what is the role of the differential returns to human capital across the two economies? Third, what is the role of differences in the stocks of human capital between these economies? Fourth, what is the role of self-selection?

I take the model to the data using a single labor force survey sampling both movers and stayers, at a point in time of a high proportion of movers. The key findings are that the "pure" effect of income differences in the choice to move to a rich economy is made up of diverse elements, operating in opposition. Thus, productivity differences in favor of the richer economy, due to differences in TFP and in the stock and quality of capital, operate to raise wages. Lower returns to human capital and lower stocks of human capital for movers, working in low-skill occupations, operate to lower wages. Self-selection on unobservables turns out to play a far smaller role. I point out how my findings relate to the results in recent literature.

It should be noted that current literature suffers from a disconcerting dichotomy. On the one hand, the development accounting literature, replete with theoretical and empirical debates, explores important income differences across economies and their determinants. On the other hand, the migration literature often looks at migration, which is driven, inter-alia, by cross-country income differences, without attempting to incorporate the insights from the former strand of literature, and without disentangling the income differences motive from a plethora of other motives. This last point results in misspecified and biased models. The empirical work in this paper addresses both of these points.

The paper proceeds as follows. Section 2 reviews the literature on rich-poor country differences and its implications for migration, and the literature on the set up of the Palestinian labor market. Section 3 presents the model. Section 4 presents the data, the econometric methodology, elaborating on the natural experiment involved, and the results. Section 5 analyzes the wage differential between movers and stayers from a number of perspectives, and includes a comparison to recent findings from other cases. Section 6 studies the patterns of self-selection on unobservables. Section 7 combines the insights from the analysis into a single consistent interpretation, offering a graphical depiction of the moving decision and a study of counterfactuals. Section 8 concludes.

2 Literature

In what follows I briefly review two main background issues: what the literature says about rich-poor country differences and their determinants and how that affects movers (sub-section 2.1) and what was the set up of the Palestinian labor market, which engendered the natural experiment (sub-section 2.2).

2.1 Differences Between Poor and Rich Countries and Their Implications for Moving

In formulating the model of worker choice whether to work in a rich economy, the paper caters for features pertaining to rich-poor economy income differences highlighted by recent literature. The development accounting literature has devoted much attention to these differences and their determinants, prominently among them, productivity differences. Recent papers have manifested disagreements with respect to the theoretical approach and its ensuing empirical implications. See, for example, the debate and discussions in Ciccone and Caselli (2019) and Jones (2019). It emerges that the theoretical formulations of output production and of human capital formation affect measurement and interpretation.

Productivity differences across countries. Jones (2016) offers a review of the evidence. Section 4.5 of his survey paper documents very substantial differences in GDP per worker across countries, with a big share, 64% on average, attributed to TFP differences. For example, his Table 6, computed using the Penn World Tables 8.0 for the year 2010, shows that advanced Western economies have 70%-80% of U.S. GDP per worker, key Latin American countries have about 35% of the U.S. value, Brazil has 18%, China 14%, and India 10%. He then reviews a number of explanations for these differences, mostly having to do with misallocation. In particular, misallocation at the micro level shows up as a reduction in total factor productivity at the aggregated level.

Acemoglu and Dell (2010) offer another direction for explanation of these income differences. They point to variation in TFP levels and in the intensity of capital use across countries (and regions) as connected to institutions. These include the enforcement of property rights, entry barriers, and freeness and fairness of elections for varying levels of government. Institutions have important implications for policy outcomes, such as the provision of public goods necessary for production and market transactions.

Human capital differences across countries. One natural idea about income differentials across countries is that they may arise from different stocks of human capital and different returns to it. The afore-mentioned literature on development accounting has found, at an earlier stage, that this latter channel does not explain much. This conclusion has recently been seriously challenged.

irJones (2014) made the point that the productivity gains associated with human capital investments cannot reveal themselves through relative wages alone, unless workers are perfect substitutes, which is an unrealistic assumption. He suggested a generalized accounting approach, allowing for imperfect substitution. This approach shows that human capital variation can account for a big part of the large income differences across countries. In Jones (2014, 2019) he computes an example of Israel, as a rich country, and Kenya, as a poor country.[3] Real GDP per capita is 16.9 times higher in Israel. The conventional approach, associated with Ciccone and Caselli (2019) for example, would posit that the human capital stock is roughly the same in the two economies and so does not explain the huge income difference. With a generalized accounting approach, if the elasticity of substitution between skilled and unskilled workers, rather than being infinite, is 1.5, then Israel has 11.7 times the human capital of Kenya.[4]

Lagakos, Moll, Porzio, Qian, and Schoellman (2018a) use representative large-sample household surveys from 18 countries with individual-level data on educational attainment, labor earnings, and the number of hours worked, and find that experience-wage profiles are on average twice as steep in rich countries as in poor countries. The same authors, Lagakos, Moll, Porzio, Qian, and Schoellman (2018b), use data on immigrants to the U.S. from the 1980–2000 US population censuses and the 2005–2013 American Community Surveys from the Integrated Public Use Microdata Series. They find that returns to experience are lower among immigrants from poor countries than among immigrants from rich countries. This holds true both for returns to "foreign experience," acquired before migrating, and returns to "US experience," acquired in the United States after migrating. These papers look at various mechanisms that can account for these differentials, including selection and loss of skill. They reach the conclusion that poor countries have lower returns, associated with lower human capital stocks, rather than the other factors (selection and skill loss). They suggest that a possible explanation would be that

> "the quantity and type of schooling result in less "learning how to learn" among individuals who attend school in poor countries. We have found support for this hypothesis by documenting that the returns to US experience among foreign-educated workers are lower than the returns to US experience for natives. At the same time, we have also documented that the returns to US experience among US educated workers are very similar to those of natives. Combined, these two facts suggest a complementarity between both quantity and type of education and subsequent human capital accumulation..."(p.334).

Hendricks and Schoellman (2018) use new data on the pre- and post-migration wages of immigrants to the United States to measure wage gains at migration. They frame their study as follows. The importance of physical capital and TFP is manifested in the wage gain at migration, relative to

[3] See Jones (2014, Tables 1 and 2, pp. 3763-3764) and Jones (2019, Table 1, p.1177).

[4] Jones (2014) reports a range of estimates for human capital ratios depending on the assumed elasticity of substitution. He finds a range between 4 and 22, using elasticities of substitution between 2 and 1.4.

the difference in GDP per worker. An immigrant has the same human capital but different physical capital and TFP before and after migration. The wage gain at migration measures the relative importance of these country-specific factors, while the residual can be attributed to gaps in human capital per worker. Their key findings are that the average immigrant from a middle-income or poor country increases their wage by a factor of 2 to 3 upon migration, which is considerably less than GDP per worker differences, which range between a factor of 6 and a factor of 32. They show that switching countries accounts for 40% of cross-country income differences while human capital accounts for 60%.

Self-Selection. Borjas, Kauppinen, and Poutvaara (2019) review the implications of the Roy model for self-selection patterns, including issues of stochastic dominance of the skill distributions of movers and stayers. They emphasize the distinction between observed and unobserved skills. Looking empirically at movers both from poor to rich economies (cited by them from key studies over the past decade) and across rich economies (their own study) they delineate the conditions for positive or negative self-selection and for stochastic dominance. In the work below, these patterns are identified for the current case and tied to the afore going discussion on productivity and human capital differences .

Implications for movers to a rich economy. There is a literature on the gains to the global economy from the move of workers from poor to rich economies; Dustmann and Preston (2019) offer a review.[5] Here I discuss two key papers.

Kennan (2013) presents a general equilibrium model, which is then evaluated empirically. He shows that if workers are much more productive in one country than in another, restrictions on immigration lead to large efficiency losses. Kennan quantifies these losses, using a set up in which efficiency differences are labor-augmenting, and free trade in product markets leads to factor price equalization, so that wages are equal across countries when measured in efficiency units of labor. The estimated gains from removing immigration restrictions are found to be large. Using data for 40 countries (see his Figure 6 and Appendix Tables 1 and 2), the average gain is estimated at $10,798 per worker per year (in 2012 dollars, adjusted for PPP), compared to average income per worker in these countries of $8,633. Thus the gain in net income is 125%. For all the countries in the Penn World Table that are not at the productivity frontier as the model defines it, using GDP data to estimate relative wages, the estimated gain is $10,135, relative

[5] In their review, Dustmann and Preston (2019) discuss a number of issues, such as goods trade, worker skill types, market structure, the roles of technology and other factors of production, taxation, and more. While the current paper studies the move from poor to rich economies, it focuses on the wage gains accrued and their decomposition, within a natural experimnt setting. It does not aim to offer the kind of wider analysis discussed in this review.

to an average income of $9,079, so the gain is 112%.

Clemens, Montenegro, and Pritchett (2019) estimate sizeable real wage gaps between migrants from 42 countries in the United States and observably equivalent workers in the origin country. These are mostly poor to rich moves. Their empirical work focuses on male workers in their late thirties, with 9–12 years of education. Their estimates indicate that for workers from the median country the relevant wage ratio (migrant to stayer) is 4.54, for the 80th percentile country is 7.58, and for the working-age population weighted average is 6.83.

In what follows, I connect the model and the empirical results to the afore-cited findings.

2.2 Background on the Palestinian Labor Market and the Natural Experiment

In the empirical work, I use Labor Force Survey (LFS) of the Israeli Central Bureau of Statistics (CBS) micro data on Palestinian males working in the local, Palestinian economy and in Israel, dating from 1987. The reason for choosing this particular time period will become clear below. As a background for the model and for the empirical methodology, I characterize Palestinian workers along some key dimensions. Angrist (1995) offers a description of the data set and of the Palestinian labor market.

The West Bank and the Gaza Strip – the constituents of the Palestinian economy – were occupied by Israel since June 1967. In 1968 Palestinian workers started to flow to employment in Israel and the labor market turned out to be the major link between the two economies.[6]The share of salaried employees employed in Israel started off at 22% in 1970, climbed to around 50% three years later, and then fluctuated around that rate and up to 65%, starting to fall off in the late 1980s. Hence, a key employment decision of the Palestinian male worker was the choice of employment location – Israel or the local economy. Men constitute the bulk of the Palestinian labor force: labor force participation rates for men aged 14 and above in the sample period were about 70%, while women had low participation rates, 7% on average.

Beginning in December 1987 the labor links between the Israeli and the Palestinian economies underwent a series of severe shocks: at the latter date a popular uprising (the first 'intifada') broke out against the occupation, leading to strikes, curfews and new security regulations, such as occasional closures of the territories. In 1993, following peace negotiations, the Oslo accords were signed, giving the Palestinians autonomous control over parts of the West Bank and the Gaza Strip. In September 2000 a second uprising

[6]Razin and Sadka (1993, Chapter 5) discuss the interdependence of the Palestinian economy and Israel, especially in terms of this export flow of labor services.

broke out, with even greater ensuing turbulence. Following the August 2005 Israeli withdrawal from the Gaza Strip there have been more violent confrontations. Consequently Palestinian employment in Israel since the end of 1987 was much more volatile and, generally, on a declining trend.[7]

In this paper I use data on Palestinian workers from 1987, a period of high Palestinian labor market involvement in Israel, pre-dating the turbulent events cited above. I elaborate more on the sample period choice and on the sample statistics below, including the education, age, industry, and occupation distributions of the Palestinian workers across the various locations. In the sample year there were no restrictions on Palestinians working in Israel nor any special screening process. The model below relates to two groups – movers and stayers; there was no other major location decision and hence no third group. Workers typically stayed a 5-6 days a week in Israel.[8]

An important fact in the present context is that there was a substantial rich-poor country difference. In the sample period, GDP per capita in the Palestinian economy was 20% of the Israeli level using data for both economies from the CBS, in local currency and current prices.[9] The World Bank puts it at 16%, for that year, using a PPP methodology. This ratio did not change much since then; the World Bank puts the average and the median at 14%, in the 24 year period from 1994 to 2017.[10]

3 The Model

The model is based on the seminal work of Roy (1951) on self-selection. The model was developed and applied empirically by Heckman and Sedlacek (1985), whose notation is followed here. As is well known, the model has been applied to labor market issues on many occasions.[11]More recently, Autor and Handel (2013) provided theoretical and empirical exploration of the relations between wages, jobs, and tasks within the framework of this model. Using job and task data, they tested the model's predictions for the relationship between tasks and wages, showing empirical support for the model. In sub-section 3.1 the basic model is presented. In sub-section 3.2 I connect insights from the recent literature, discussed above, to the various components of the model.

[7]For details on developments in the Palestinian labor market, see Bartram (1998). For an analysis of the Israeli labor market, see Yashiv (2000).

[8]Semyonov and Lewin-Epstein (1987) COMPLETE. See Dustmann and Gorlach (2016) for a discussion of migrants whose duration in the host country is limited.

[9]Source: Tables 2.1, 6.7, 27.1 and 27.9 in the 1991 CBS Statistical Abstract.

[10]Computation is in PPP terms; See https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=IL-PS

[11]See, for example, Borjas, Kauppinen, and Poutvaara (2019) and references therein, for the migration context.

3.1 The Movers Decisions

Tasks and production. There are two localities, a rich and a poor one indexed $i(= P, R)$ in which workers can work. In the current context these are the rich, host country, Israel, and the poor, source economy, the Palestinian one. Workers are free to enter the economy that gives them the highest income but are limited to work in only one location at a time. Each location requires a unique, specific task T_i. Each worker is endowed with a vector of skills $(\mathbf{S})$ which enables him to perform location-specific tasks. The vector $\mathbf{S}$ is continuously distributed with density $g(\mathbf{S} \mid \Theta)$ where Θ is a vector of parameters. $t_i(\mathbf{S})$ is a non-negative function that expresses the amount of task a worker with the given skill endowment $\mathbf{S}$ can perform and is continuously differentiable in $\mathbf{S}$.

Aggregating the micro supply of task to location i yields:

$$T_i = \int t_i(\mathbf{S}) g(\mathbf{S} \mid \Theta) d\mathbf{S} \tag{1}$$

The output of location i is given by:

$$Y_i = F^i(T_i, \mathbf{I}_i) \tag{2}$$

where $\mathbf{I}$ is a vector of non-labor inputs. The production function F is assumed to be twice continuously differentiable and strictly concave in all its arguments. For a given output price P_i, the equilibrium price of task i equals the value of the marginal product of a unit of the task in location i. This task price will be denoted by Π_i in nominal terms and π_i in real terms:

$$\Pi_i = P_i \frac{\partial F^i}{\partial T_i} \tag{3}$$

$$\pi_i = \frac{\partial F^i}{\partial T_i} \tag{4}$$

Assuming workers are paid their marginal products, real wages per worker in this set-up are given by:

$$\ln w_i(\mathbf{S}) = \ln \pi_i + \ln t_i(\mathbf{S}) \tag{5}$$

Functional forms. I shall be using the following functional form for the task function:

$$\ln t_i(\mathbf{S}) = \beta_0 + \sum_j \beta_{j,i} S_j + u_i \tag{6}$$

Hence:

$$\begin{aligned} \ln w_i(\mathbf{S}) &= \ln \pi_i + \ln t_i(\mathbf{S}) \\ &= \ln \pi_i + \beta_0 + \sum_k \beta_{k,i} S_k + u_i \end{aligned} \tag{7}$$

Travel costs. Additionally, I postulate – to make the model consistent with the data to be examined – that the individual worker has travel costs to work. These depend on a vector of variables related to location, to be denoted **L**, and are formulated as a fraction $k_i(\mathbf{L})$ of wages:

$$\text{travel costs} = k_i(\mathbf{L}) w_i \tag{8}$$

I discuss these variables in the empirical work below.

Income maximization. An income-maximizing individual chooses the location i that satisfies:

$$w_i(1 - k_i(\mathbf{L})) > w_j(1 - k_j(\mathbf{L})) \tag{9}$$

Hence:

$$[\pi_i t_i(S)] \, [1 - k_i(\mathbf{L})] > \left[\pi_j t_j(S)\right] \left[1 - k_j(\mathbf{L})\right] \quad i \neq j; \tag{10}$$

Density of Skills. Further analysis requires the adoption of specific functional forms for the density of skills g and the function mapping skills to tasks t. Roy (1951) assumed that these are such that the tasks are log-normal i.e. $(\ln t_i, \ln t_j)$ have a mean (μ_i, μ_j) and co-variance matrix Σ (with elements denoted by σ_{ij}) . Denoting a zero-mean, normal vector by (u_i, u_j) the workers choose between two wages:

$$\begin{aligned} \ln w_i &= \ln \pi_i + \mu_i + u_i \\ \ln w_j &= \ln \pi_j + \mu_j + u_j \end{aligned} \tag{11}$$

If $\ln w_1 + \ln [1 - k_i(\mathbf{L})] > \ln w_j + \ln \left[1 - k_j(\mathbf{L})\right]$, the worker chooses location i. If the converse is true, the worker chooses location j.

With these functional specifications, the following holds true:[12]

$$pr(i) = P\left(\ln w_i + \ln [1 - k_i(\mathbf{L})] > \ln w_j + \ln \left[1 - k_j(\mathbf{L})\right]\right) = \Phi(c_i)$$

where

[12]The following equations are based on the properties of incidentally truncated bivariate normal distributions.

$$\begin{aligned} c_i &= \frac{\ln\frac{\pi_i}{\pi_j} + \ln\frac{[1-k_i(\mathbf{L})]}{[1-k_j(\mathbf{L})]} + \mu_i - \mu_j}{\sigma^*}, \; i \neq j \\ \sigma^* &= \sqrt{var(u_i - u_j)} \end{aligned}$$

$\Phi(\cdot)$ the cdf of a standard normal variable. The proportion of workers in location i will increase as the task price π_i there rises, as relative travel costs for the location $k_i(\mathbf{L})$ decline, or as the mean of the task μ_i rises. In addition it depends on the variance and co-variance terms in Σ via σ^*.

3.2 Insights for Model Components from the Literature

I connect the afore-going model to the development accounting literature, discussed in sub-section 2.1 above. Note a crucial distinction with respect to this literature. In the current paper, $\ln w_i$ always refers to a wage of a Palestinian worker, not an Israeli worker, and the index i refers to the location – Israel or the local economy. Hence wage gains are going to be empirically examined across locations (of Palestinian workers), i.e., movers and stayers, not across workers of different economies (Israelis and Palestinians).

I use the framework presented in Hendricks and Schoellman (2018). In formal terms, assume a Cobb Douglas production function, with physical capital K, human capital T, and technology A:

$$Y_i = K_i^{\alpha}(A_i T_i)^{1-\alpha} \tag{12}$$

Define:

$$z_i \equiv \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} A_i \tag{13}$$

The Appendix shows that this definition and the relation $T_i = L_i t_i$, where L is the number of workers, imply that GDP per worker in logs is given by:

$$\ln\frac{Y_i}{L_i} = \ln z_i + \ln t_i \tag{14}$$

Assuming workers are paid their marginal products, real wages per worker in this set-up are given by:

$$\begin{aligned} \ln w_i &= \ln(1-\alpha) + \ln\frac{Y_i}{L_i} \\ &= \ln(1-\alpha) + \ln z_i + \ln t_i \end{aligned} \tag{15}$$

Using (7) this means:

$$\ln \pi_i = \ln(1-\alpha) + \ln z_i \tag{16}$$

Workers can gain by a move to a richer economy with a higher level of z_i (and therefore also $\frac{Y_i}{L_i}$). The worker gains because of work in an economy, with higher levels of K and/or A, as seen in equation (13). In terms of the preceding analysis, this means that the richer economy has a higher level of π_i (see equation (16)). These, however, are not the only consequences for wages. Equation (7) shows that the term $\sum\limits_j \beta_{j,i} S_j + u_i$ will be important for wages too.

I turn to discuss the findings in the literature concerning these different elements.

Productivity. The issues debated in the cited literature include, inter-alia, the size of z and T across countries.A fairly accepted finding is that z is higher in the rich economy relative to the poor one. Recent papers, reviewed in sub-section 2.1, argue also for a relatively big role of T. The debates pertain to the size of the z differential across countries, i.e., the breakdown of equation (14) in cross-country differences into components. This differential will be pertinent here, too, across locations.

The ideas of different institutions and misallocation discussed in sub-section 2.1 certainly seem pertinent to the economies in question here.

Human capital: stocks and returns. In the empirical work I posit:

$$\overline{\ln w_i} = \widehat{k}_i + \widehat{\boldsymbol{\beta}}_i \overline{\mathbf{X}}_i + \widehat{\rho}_i \widehat{\sqrt{\sigma_{ii}}} \widehat{\overline{\lambda(c_i)}}$$

and estimate the hatted variables. In terms of the literature cited above, the findings are that movers "take" both $\widehat{\boldsymbol{\beta}}_i$ and $\overline{\mathbf{X}}$ "with them",i.e., they are embodied in the workers. See, for example, the evidence presented in Lagakos, Moll, Porzio, Qian, and Schoellman (2018b), discussed above, in particular their summary on page 333. The discussion in Hendricks and Schoellman (2018) on skill transfer (see pages 692-697) leads to the conclusion that in the present case there is full skill transfer, as movers worked in their original occupations and were asked to perform similar tasks in the host, rich economy as in their home, and poor, economy.

Note, too, that T differentials across economies are a subject studied in the literature, but that here the analysis will pertain to T differentials across locations.[13]

Selection. Selection enters here via equations (10)-(**??**). The model allows for the analysis – separately – of selection on observables and on un-

[13] Note the distinction between the differential $\widetilde{T_{Israel}} - \widetilde{T_{Palestine}}$, examined in the development accounting literature, and the differential $\widetilde{T_{P\_in\_Israel}} - \widetilde{T_{P\_local}}$ (with P denoting Palestinian workers) studied here. This distinction is borne out by the findings of the recent literature.

observables. The effect of $\widehat{\boldsymbol{\beta}}_i\overline{\mathbf{X}}_i$ will express selection on observables (on the $\overline{\mathbf{X}}_i$) and the effect of $\widehat{\rho}_i\widehat{\sqrt{\sigma_{ii}}}\widehat{\lambda(c_i)}$ will express selection on unobservables.

The upshot is that there are opposing effects at work here: moving from a poor to a rich economy may increase wages via the productivity channel (expressed in z_i or π_i); human capital effects ($\widehat{\boldsymbol{\beta}}_i\overline{\mathbf{X}}_i$) could mitigate the rise in wages, if movers work in low-skill, low-return occupations. Moreover, self-selection on unobservables may work in diverse ways, positive and negative, depending upon the second moments of the u_i terms of the task function distribution.

4 Data, Methodology, and Results

In this section I estimate selection and wage equations for Palestinian men working in Israel and East Jerusalem as one location and working locally (in the West Bank and Gaza) as the other location. In what follows I discuss the data (4.1), the natural experiment (4.2), the econometric methodology (4.3), and identification and specification issues (4.4). I then report the results (4.5). The analysis and interpretation of the results are elaborated in the subsequent sections.

4.1 The Data

The data are taken from the Palestinian Territories Labor Force Survey (TLFS) conducted by the Israeli Central Bureau of Statistics (CBS); for detailed descriptions of this data set, see CBS (1996) and Angrist (1995).[14]Its principles are similar to the Israeli Labor Force Survey undertaken by the CBS, which is akin to other such surveys, such as the U.S. Current Population Survey. The survey used a 1967 CBS-conducted Census as the sampling frame, with a major update in 1987. It was conducted quarterly and included 6,500 households in the West Bank and 2,000 in Gaza, surveyed by local Palestinian enumerators employed by the Israeli Civil Administration in the Territories. The TLFS sampling frame includes most households in the West Bank and Gaza Strip, regardless of the employment status or work location of the head of household. It included questions on demographics, schooling and labor market experience.

In this paper I use observations on Palestinian men[15] aged 18-64 from the TLFS in the year 1987. The reason for the choice of this year is that it represents the time of highest data quality, following the sample frame revision, and, as mentioned, a high share of Palestinian employment in Israel. It was the last one before the uprising and the ensuing turbulence.

[14]I am grateful to Joshua Angrist for the use of his processed version of the TLFS data set.

[15]As mentioned, women had very low participation rates, and when working in the market economy, did so locally, not in Israel.

It should be noted that running the same analysis over cross-sections from this survey in the years 1980-1986 yielded very similar results.

Table 1 presents sample statistics for the variables used in the empirical analysis.

Table 1

The table shows that, on average, local workers (stayers) earned lower wages and were more educated and more experienced than workers in Israel (movers). Average schooling levels are consistent with the features of a developing economy. Decomposing each group into types of residence, it can be seen that rural residence was the main type for movers. For stayers, rural and urban residence had similar employment shares. I provide further information on the employment characteristics (industries and occupations) of these workers and on worker skill levels, when discussing the relevant estimation results below.

4.2 The Natural Experiment

The set up described above precludes confounding with other factors driving movers, which I delineate as follows. These factors have been extensively cited in the literature; for a recent analysis see Dao et al (2018).

Geographical distance. The distance to be travelled is an obvious determinant, affecting costs, including possibly socio-psychological costs. In the current case this distance was travelled, usually weekly, in a matter of 30 to 90 minutes. Hence, while it can be used to facilitate identification as done below, it did not generate large scale costs.

Family linkages and local social networks. Movers may be motivated by the wish to join families in host economies or by the possibility to use local migrant networks. This is not the case here, as the families of movers did not leave their homes; work was done by daily or weekly commute; and there was no host economy network.

Credit constraints. Credit constraints may play a big role in moving decisions. The costs involved may be such that they require taking out loans. In the current case, costs were relatively small. In many cases the relevant costs, such as transportation and housing, were paid for by the employers, albeit partly out of wages. This did not necessitate the use of loans.

Welfare benefits. Movers are frequently attracted by the possibility to receive welfare benefits and various other forms of social assistance from host economies. This was completely absent in the current case.

Insurance motives. Movers may be concerned in some cases with negative events or shocks in the home economy, actual or anticipated. Moving has therefore a kind of insurance motive, including from the perspective of the wider family. This kind of motive played a certain role after 1987, when adverse shocks did occur. But in the sample period this kind of motive did not exist.

Social-Psychological issues. Movers are often affected by difficulties in leaving home for social and psychological reasons. In this case the separation from home was very short-lived, a few days in succession at the most. Hence this determinant had much less power.

In formal terms, the empirical formulation of the movers' problem is generally given by:

$$M = f(w_i - w_j, \mathbf{X}) + \varepsilon$$

where M is the moving decision, $w_i - w_j$ are wage differences, $\mathbf{X}$ is a vector of determinants such as the one discussed above, and ε is a random effect. Equation (9) above is a special case. In the current case of Palestinian workers there were virtually no elements in the vector $\mathbf{X}$. But in most cases this does not hold true, i.e., the vector $\mathbf{X}$ is not empty, but nonetheless the model is estimated, without at least some of the elements of $\mathbf{X}$.

4.3 Econometric Methodology

Estimation of equations (11) for workers employed locally and employed in Israel will yield estimates of all the key elements of the model, i.e., $\ln \pi_i$, μ_i and the elements of Σ. To do that the following procedure is used:

(i) I posit that $\ln t_i = c_i S$ where S is decomposed into observed and unobserved variables S_o and S_u, and c_i their associated coefficients, are c_{io} and c_{iu}, respectively. Thus equations (11) become:

$$\ln w_i = \ln \pi_i + \beta_i \mathbf{X} + u_i, \tag{17}$$

where $\beta_i = c_{io}, \mathbf{X} = S_o$ and $c_{iu} S_u = u_{i.}$

(ii) When estimating equation (17), I take into account sample selection, which is inherent in the model. Thus define the variable z^* :

$$\begin{aligned} z^* &= \ln w_i + \ln(1 - k_i(\mathbf{L})) - \ln w_j - \ln(1 - k_j(\mathbf{L})) \qquad (18) \\ &= \ln \pi_i - \ln \pi_j + \ln(1 - k_i(\mathbf{L})) - \ln(1 - k_j(\mathbf{L})) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j \end{aligned}$$

and the indicator variable z :

$$\begin{aligned} z &= 1 \; if \; z^* > 0 \qquad (19) \\ z &= 0 \text{ otherwise} \end{aligned}$$

According to the model one observes $\ln w_i$ only if $z^* > 0$ i.e., when $z = 1$. Paralleling (**??**) we have:

$$\Pr(z = 1) = \Phi(\ln \frac{\pi_i}{\pi_j} + \ln \frac{(1 - k_i(\mathbf{L}))}{(1 - k_j(\mathbf{L}))} + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \tag{20}$$
$$\Pr(z = 0) = 1 - \Phi(\ln \frac{\pi_i}{\pi_j} + \ln \frac{(1 - k_i(\mathbf{L}))}{(1 - k_j(\mathbf{L}))} + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j)$$

The observed $\ln w_i$ is given by:

$$\ln w_i \mid (z = 1) = \ln \pi_i + \beta_i \mathbf{X} + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}\right] \lambda(c_i) + u_i \tag{21}$$

where:

$$c_i = \frac{\ln \frac{\pi_i}{\pi_j} + \ln \frac{[1 - k_i(\mathbf{L})]}{[1 - k_j(\mathbf{L})]} + \mu_i - \mu_j}{\sigma^*}, \; i \neq j$$
$$\lambda(c_i) = \frac{\phi(c_i)}{\Phi(c_i)}$$
$$\sigma^* = \sqrt{var(u_i - u_j)}$$
$$\rho_i = correl(u_i, u_i - u_j), \; i \neq j; i, j = 1, 2$$

with $\phi(\cdot)$ denoting the density of a standard normal variable.

This may also be written as follows:

$$\ln w_i \mid (z = 1) = \ln \pi_i + \beta_i \mathbf{X} + \rho_i \sqrt{\sigma_{ii}} \lambda(c_i) + u_i \tag{22}$$

A similar equation holds true for the other location. Note that while the **X** vector appears in both (20) and (22), the **L** vector appears only in the selection equation (20). I estimate the model using Heckman's (1979) two-step consistent estimation procedure. One can interpret the selection bias in (17) as an omitted variable bias. If $\lambda(c_i)$ is not included in the equation, the estimates of the vector of coefficients β_i may be biased. The sign of the bias depends on the effect of x_k on selection and on the effect of selectivity on the dependent variable, i.e., on wages in this case. The following equation expresses this bias formally. For any variable x_k in **X**:

$$\frac{\partial E(\ln w_i \mid (z = 1))}{\partial x_k} = \beta_{ik} + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}\right] \frac{\partial \lambda}{\partial c_i} \frac{\partial c_i}{\partial x_k} \tag{23}$$

The sign of the bias depends on the type of selection process ($\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}$) and on the direction of influence of the relevant variable on the locational selection ($\frac{\partial c_i}{\partial x_k}$). The magnitude depends on these factors as well as on the $\frac{\partial \lambda}{\partial c_i}$ term.

4.4 Identification and Specification

The identification problems of selection models have been much explored and are well-known. The way the model here can be estimated using exclusion restrictions is by postulating variables that affect travel costs, and hence selection, but not wages.[16]There is one variable that clearly fits this requirement – geographical regions or localities. This is a useful measure of the determinants of travel costs because workers' homes are located in different distances from the locations of employers.

Two other variables are "candidates" but may arguably be affecting wages too, and so are weaker as exclusion restrictions: one is the type of residence. This variable includes rural areas, urban areas, and refugee camps. These may serve to indicate travel costs as rural residents are likely to be more spread out and refugee camps residents are likely to be more concentrated. In camps there are likely to be organized, common means of transport. The other candidate variable is marital status. This variable is not directly related to travel costs but may serve to indicate costs that pertain to the economic life of the household.

The data sample does not contain other variables relating to the household which could provide additional exclusion restrictions. I therefore use the geographical variable as the sole restriction in the benchmark case. Additionally, I use the above two variables as a variation on the restrictions, albeit these not being ideal choices for instruments.

For the travel cost function $k_i(\mathbf{L})$, included in the selection equation only, I postulate the following:

$$k_i(\mathbf{L}) = \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^i$$

where l is the region of the worker's residence, p is an index of regions, θ_p is a coefficient to be estimated; the Y_n variables are additional variables affecting travel costs and γ_n are their coefficients to be estimated; as before, location i indicates the local or host economy. The θs and the γs are estimated in the selection equations (20). The l_p variables are the dummy variables for geographical regions or localities discussed above. The Y_n variables are the type of residence and marital status variables. Summary statistics of these variables appear in Table 1 above.

For the task function variables $\mathbf{X}$, included in both the selection and wage equations, I use education and a linear-quadratic formulation for experience[17]. I also use indicator variables for the quarters within 1987, which I do not report.

Approximating I get:

[16]For a recent discussion of the use of exclusion restrictions see Wooldridge (2015),

[17]Experience being defined as age minus education minus 5.

$$
\begin{aligned}
\ln(1-k_i(\mathbf{L})) &= \ln(1-\sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^i) \\
&\simeq -\sum_p \theta_p \cdot l_p^i - \sum_n \gamma_n Y_n^i
\end{aligned}
$$

The selection equations are thus:

$$
\begin{aligned}
\Pr(z &= 1) = \Phi(\ln \frac{\pi_i}{\pi_j} + \sum_p \theta_p \cdot l_p^j - \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^j - \sum_n \gamma_n Y_n^i + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \quad (24) \\
\Pr(z &= 0) = 1 - \Phi(\ln \frac{\pi_i}{\pi_j} + \sum_p \theta_p \cdot l_p^j - \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^j - \sum_n \gamma_n Y_n^i + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j)
\end{aligned}
$$

The estimated wage equation is the following:

$$
\begin{aligned}
\ln w_i \quad | \quad & \text{location } i = \ln \pi_i + \beta_{0i} + \beta_{1\,i} educ + \beta_{2\,i} exp + \beta_{3i} exp^2 \\
& + \sum_{m=2}^{4} \gamma_m Q_m + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*} \right] \lambda\left(c_i\right) + u_i \qquad (25)
\end{aligned}
$$

where i, j denote locations, Q is an indicator variable for the quarter, and m denotes the quarter number. The dependent variable in the wage equation is the log of real hourly wages ($\ln w_i$), defined as the nominal monthly wage divided by hours worked and deflated by the CPI.[18] The use of hourly wages is designed to avoid confounding the choice of work place with the choice of work time (hours or days).[19] Education ($educ$) and experience (exp) are defined in years.

The benchmark specification reported below [column (1) of Tables 2 and 3] has the geographical exclusion restrictions.. The alternative, specification 2 includes the variables discussed above contained in **L**, so there are three exclusion restrictions. Specification (3) uses OLS to test for the effect of selection correction (running only the wage equation).

4.5 Results

Tables 2 and 3 report the results. Table 2 reports the estimates of the selection equation and Table 3 reports the estimates of the wage equation for

[18]Real, rather than nominal, wages are used in accordance with the model presented in Section 3. Moreover, inflation was relatively high (16.1%) in the course of 1987. The same CPI applies to both locations.

[19]I delete observations of nominal hourly wages less than 0.1 NIS and higher than 11.5 NIS. These are the lowest 1% and highest 0.2% of the wage distribution. For these observations wages are either extremely low or unreasonably high, indicating that they are either measured with error or that they reflect very few hours of monthly work. A similar procedure was employed by Heckman and Sedlacek (1985).

the specifications discussed above. In each case I report the point estimates with standard errors in parentheses; in the wage regressions I also report the implied second moments (ρ_i, σ_{ii} and ρ_{ij}), and the Wald test (using χ^2 test statistics, with p-values in parentheses).

Tables 2 and 3

Consider first the differences across specifications:

OLS vs. selection-corrected Heckman estimates. The OLS estimates imply smaller differences between the two economies in all estimated parameters. I expand on this issue below.

The effect of the exclusion restrictions. The emerging picture across columns 1 and 2 is the same, but there are two differences: column 1 has higher point estimates in absolute value for the four estimated parameters and it implies a lower correlation of the unobserved skills distributions relative to column 2.

The following summarizes the results in terms of *location selection and the effects on wages*:

(i) Selection to move is negatively related to education, experience, refugee camp and urban residence, and is positively related to being married.

(ii) The constant of the wage equation is substantially higher in Israel.

(iii) Education and experience premia are higher in local employment relative to employment in Israel. Consistent with this finding are the afore-cited selection equation results, whereby education and experience decrease the probability of choosing employment in Israel.

(iv) Estimates of the second moments indicate higher variance of the local unobserved skills distribution. The ratio of the standard deviations $\frac{\sqrt{\sigma_{israel}}}{\sqrt{\sigma_{local}}}$ is around 0.8.

(v) The correlation between the unobserved skill distributions is lower than the ratio of standard deviations, i.e., $\rho_{israel,local} < \frac{\sqrt{\sigma_{israel}}}{\sqrt{\sigma_{local}}}$, and is around 0.6-0.7.

I turn now to examine the implications of these results. In what follows I use the results of column 1 of Tables 2 and 3 as the benchmark results (noting that column 2 delivers very similar implications).

5 The Wage Differential Between Movers and Stayers

I decompose the wage differential and study its components. Note that the wage differential explored here is the one between Palestinian workers movers and stayers, not between native workers of the two economies. I analyze it from a number of aspects.

5.1 Components of the Wage Differential

In Table 4 I quantify the relative role played by the different elements of the model – task prices, skill premia, skill levels, and selectivity effects. I do so using actual data and the point estimates of column (1) from Table 3.

Table 4

Panel (a) of Table 4 reports the elements of mean wages in each of the locations, using the following equations:

$$\begin{aligned}\overline{\ln w_{local}} &= \widehat{k}_{local} + \widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} + \widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} \\ \overline{\ln w_{Israel}} &= \widehat{k}_{Israel} + \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel} + \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}\end{aligned} \tag{26}$$

where $\overline{\ln w_i}$ is the mean log hourly wage in economy $i, \widehat{k}_i = \ln \widehat{\pi_0} + \widehat{\beta}_0$ for economy i using the point estimates of the wage equation's constant, $\widehat{\boldsymbol{\beta}}_i$ is a vector of the point estimates of the coefficients in economy i, $\overline{\mathbf{X}}_i$ is a vector of the mean values of the independent variables in economy i, and $\widehat{\rho}_i\widehat{\sqrt{\sigma_{ii}}}\widehat{\overline{\lambda}_i}$ are the estimates of the second moments times the average of the estimated inverse of Mills' ratio.

Panels (b) and (c) of Table 4 report the mean wage differential between movers and stayers. I decompose the mean wage differential between Palestinian workers in the Israeli economy and in the local economy into components: a part due to task prices plus the intercept of the task function (i.e., the constant in the wage equation); a part due to differences in skill premia across the two locations; a part due to differences in skill levels across the two locations; and a part due to differences in selection effects. This is done in two alternative ways (panels b and c), elaborated in the table.

The key findings from the table are as follows.

Moving premium. The wage equation's intercept – reflecting the task price π_i and the constant term in the task function – is substantially higher in Israel. Note that this difference in baseline wages, or 'moving premium,' is much higher than the difference in mean wages between Israel and local employment: the difference in the constant of the equation between Israel and local employment is 0.71 log points while the difference in average wages is 0.09 log points (both in terms of the log real hourly wage). This big difference shows that there is a large offset to the moving premium, which in itself is large. I discuss this offset in detail, below.

Skill premia.

(i) Locally, the schooling premium is above 4%; in Israel it is estimated to be about 1%.[20]

[20]The very low returns to schooling for Palestinian men in the Israeli economy are consistent with the findings of Angrist (1995, Table 4).

(ii) Locally the experience premia profile of earnings has the familiar hump-shape while in Israel it is relatively flat and low.

Selection on Observables. Less educated and less experienced workers chose to work in Israel; those with better skills chose to work locally and were compensated for the baseline wage differential by the local returns given to their skills. This represents negative selection on observed skills. Borjas, Kauppinen, and Poutvaara (2019) show[21] that the skill distribution for stayers stochastically dominates the distribution for movers in this case, whereby the rate of return to observable skills is higher at home.

This sorting pattern, implied by the results of estimation, is borne out by the actual, observed locational distributions by education and age. Table 1 above has presented key moments for education and experience. The following table offers additional evidence by describing the distribution of workers across work locations by education and age:

Table 5

The table confirms that it is indeed the less educated and younger workers who worked relatively more in Israel. Locally, mean schooling and age are higher. Particularly striking are the results for the high schooling group, where the share of workers is substantially higher in local employment.

Tasks, skill premia, and selection. How can one account for the fact that the returns to the same skills differ markedly for movers and stayers? The local economy rewarded education and experience substantially more, which can be explained by looking more closely at the types of jobs in each economy. Table 6 shows the distribution of employment across industries and occupations.

Table 6

Local employment was characterized by industries and occupations that presumably require the performance of more analytical tasks. In particular, government, personal, and financial services are about 40% of local employment. In contrast, in Israel employment was highly concentrated (over 80%) in three industries – construction, manufacturing and agriculture, typically requiring manual tasks. In terms of occupations, 21% of local workers were employed in high-skilled occupations (the top three in the table) vs. 2% in such occupations in Israel. Hence it is not surprising that local employment offered higher returns for education and experience. This set-up is consistent with the formulations of the model, whereby the two locations require the performance of different tasks T_i and which rewards skills differentially. This pattern is consistent with the findings of Autor and Handel (2013) using detailed task and job data.

[21]See their page 150 and equation 12.

This phenomenon of low skill premia for movers is also consistent with the recent findings in the literature, pertaining to movers from poor to rich economies, cited and discussed in sub-section 2.1 above. Moreover, the low returns to experience are consistent with the results of Dustmann and Meghir (2005), who studied returns to experience for young German workers. They found that much of the return to low skilled workers is due to such workers finding good matches and remaining with them. The case of low skilled Palestinians in Israel is likely to violate both requirements – there is no search process for good matches and the employment relationship is not of long duration.

5.2 Comparison to Recent Findings

In the following computations I derive implications of these estimates, which gives a sense of where these results stand with respect to findings in recent literature.

5.2.1 Wage Gaps Between Movers and Stayers

Clemens, Montenegro, and Pritchett (2019) estimate real wage gaps between migrants in the United States and their observably-equivalent national counterparts in 42 home labor markets. They use theory to bound a wage gap called a 'place premium' as it does not arise from portable individual traits. In the analysis here I obtain estimates of such gaps.

Clemens et al (2019) define the following log wage ratios, which can be represented as follows in terms of the current model:

$$\begin{aligned} \ln R_u &= \ln \pi_1 - \ln \pi_2 + \mu_1(E_1 s) - \mu_2(E_2 s) + E_1 u_1 - E_2 u_2 \\ \ln R_c &= \ln \pi_1 - \ln \pi_2 + \mu_1(E_1 s) - \mu_2(E_1 s) + E_1 u_1 - E_2 u_2 \end{aligned} \tag{27}$$

where R_u is the unconditional ratio of migrants' mean wages in the host economy to mean wages in the home economy, without adjustment for observable or unobservable differences between average migrants and average non-migrants. R_c is the ratio conditional on observable inherent differences like age and education. The decomposition reported in Table 4 above is related to these ratios (27) as follows:

$$\begin{aligned} \ln R_u &= \overline{\ln w_{Israel}} - \overline{\ln w_{local}} \\ \ln R_c &= \widehat{k}_{Israel} - \widehat{k}_{local} + \overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel}) + \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}} - \widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} \end{aligned} \tag{28}$$

Table 4 provides estimates of the constituents of these ratios. Given the numbers in the table I get:

$$
\begin{aligned}
-\ln R_c &= -0.71 + 0.54 + 0.02 = -0.15 \\
\ln R_u &= 0.09
\end{aligned}
$$

Hence $\ln R_u - \ln R_c = 0.09 - 0.15 = -0.06$ and $\frac{R_u}{R_c} = 0.94$. This figure is very close to the numbers reported by Clemens et al (2019, Table 3) relying on four studies of the migration from Mexico, Puerto Rico and Nicaragua to the U.S. which range from 0.85 to 0.89, as well as one study on migration from Romania to Spain, which finds a value of 0.87.

5.2.2 Wages, Technology and Physical Capital, and the Role of Human Capital

Following the derivation discussed in sub-section 3.2 and in the Appendix, and using equation (16), productivity differences across locations are given by:

$$\ln z_i - \ln z_j = \ln \pi_i - \ln \pi_j \tag{29}$$

The estimates of Table 4 imply that this difference is 0.71 log points in favor of the Israeli economy or a $\frac{z_{Israel}}{z_{local}}$ ratio of 2. 03. Hendricks and Schoellman (2018) report z ratios of similar magnitude.[22]Note that the total wage gain of 0.09 log points masks this substantial gain in productivity. The masking is due to the offsetting effect of skills and their returns.

In this context, one should note important, but subtle, differences between the current analysis and the development accounting framework.

The latter framework underlies the analysis of Hendricks and Schoellman (2018), henceforth HS. The HS computation is based on data – GDP per capita and pre- and post- migration wages. It assumes, in the baseline scenario, that human capital is fully transferable. It is thus able to deduce the country effect, related to the levels of its technology and physical capital, by comparing log differences in GDP per capita to log differences in the afore-cited wages, across the U.S. and source countries (see their equation 4). HS find that wage differentials are lower than GDP per capita differentials (their Table II). Wage gains vary by factors ranging from 1.2 to 3.2, while GDP per capita gaps vary by factors ranging from 1.3 to 31.8. Hence, HS reach the conclusion that human capital differences play a big role, between 48% and 66%, in cross country income differences. In their conclusions (page 697) they use the number 62% to describe the case that applies to the standard development accounting framework. This conclusion remains true when taking into account imperfect substitutability, selection,

[22]See in their Table II, the column of 'wage gain,' which, by their equation 4, captures the log difference in z. It ranges from 1.2 to 3.2.

and skill transfer effects (see their Tables IV,V and VIII). The contribution of HS to the development accounting literature is to flag the role of human capital differences.

The current paper has tasks, rather than human capital stocks per se, in production (see equation (2) above). Tasks are defined by location and are bundles of skills, with returns to these skills included (see equation (6) above). Hence workers are paid according to the relevant task bundle in a given location. When comparing locations, the z's (technology cum physical capital, see equation (13)) of a location reflect the country. The task bundle reflects the worker (his skills, $\mathbf{X}$) and the task returns ($\boldsymbol{\beta}$). The wage differential here, across locations, is thus not the same as the HS wage differentials. The wage differential here reflects both the z cross-country differential (as in HS) **as well as** the task differential across locations, which reflects worker skills and task returns (unlike HS).

The HS results are likely to hold true in the current case. It is highly likely that human capital is higher in Israel and that it plays a big role in the GDP per capita differential (which is a factor of about 5 here). These points, however, are not examined in the current paper. Likewise, the findings, whereby the foreign task bundle has low value in terms of wages for the movers, is not an issue examined by HS. This low value is consistent with both the HS view on lower human capital in the poor country, and the findings, related to poor countries human capital, of Clemens et al (2019) and Lagakos et al (2018a,b), cited above. Thus, large differences in human capital explain the offset effect here, through task values, which lowers the wages of movers. This is an important implication of the analysis, given that the findings on human capital differences across countries have been highlighted in the litearture only recently.

6 Self-Selection on Unobservables

The preceding section has decomposed the wage differential between movers and stayers. Part of this differential is due to selection on observables as discussed above. In this section I study the estimated selection patterns related to unobserved skills.

6.1 Patterns of Self-Selection

Post-selection the *conditional* mean and variance of the locational wage distribution can be characterized; note that these will also characterize the *observed* distribution if the model holds true:

$$E\left(\ln w_i \mid \ln w_i + \ln\left[1-k_i(\mathbf{L})\right] > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right]\right) = \ln \pi_i + \mu_i + \frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}\lambda(c_i) \quad (30)$$
$$var\left(\ln w_i \mid \ln w_i + \ln\left[1-k_i(\mathbf{L})\right] > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right]\right) = \sigma_{ii}\left\{ \begin{array}{c} \rho_i^2[1-c_i\lambda(c_i)-\lambda^2(c_i)] \\ +(1-\rho_i^2) \end{array} \right\}$$

It is possible to classify the selection outcomes in terms of the relations between the elements of Σ: σ_{ii}, σ_{jj} and σ_{ij} or alternatively between $\frac{\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}}}$ and $\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$.[23] Assuming, without loss of generality, that $\sigma_{jj} \geq \sigma_{ii}$, the different outcomes depend on the relation between the ratio of the standard deviation in each location $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ and the correlation between the two locational distributions ρ_{ij}.

Three cases are possible:[24]

(i) The correlation between the countries is positive and relatively high, i.e., $\rho_{ij} \geq \frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$. In this case the term $\frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}$ in equation (30) is positive for location *j* and negative for location *i*. Thus the conditional mean in location *j* (location *i*) i.e., the mean expressed by equation (30), is higher (lower) than the unconditional mean, $\ln \pi_i + \mu_i$ (note that $\lambda(c_i)$ is positive). Selection is positive in location *j* and negative in *i*. Because of the high correlation, this is a comparative advantage case rather than absolute advantage, i.e., workers who do well in a certain location may still select the other one and workers may select a location that they do badly in.

(ii) The correlation between the countries is negative, i.e., $\rho_{ij} < 0$. In this case the term $\frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}$ in equation (30) is positive for each location so the conditional mean in each location is higher than the unconditional mean. This is a case of positive selection in the two countries or of absolute advantage – each location tends to be filled with the workers that perform best in the location.

(iii) The correlation between the countries is positive but relatively low, i.e., $0 \leq \rho_{ij} < \frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$. In this case too the term $\frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}$ in equation (30) is positive for both countries, and in each location there is positive selection, though it

[23]Note the following definitions which will appear below:

$$\rho_1 = \frac{\sigma_{ii}-\sigma_{ij}}{\sqrt{\sigma_{ii}}\sigma^*}$$
$$\rho_2 = \frac{\sigma_{jj}-\sigma_{ij}}{\sqrt{\sigma_{jj}}\sigma^*}$$
$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$$

[24]Remarking that ρ_{ij} is bounded from above by $1 \leq \frac{\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}}}$.

is once more comparative and not absolute advantage which dictates selection. Note that this case includes $\rho_{ij} = 0$, i.e., the endowment of tasks are uncorrelated.

Note that task prices and mean abilities operate through c and $\lambda(c)$. They do not determine the afore cited selection patterns but they do affect the magnitude of selection.

6.2 Implications of the Estimates

Selection bias. How does self-selection affect the estimates and how does it affect mean wages? Table 3 above has reported the education and experience coefficients for the wage equations using OLS not corrected for sample selection bias (column 3), which can be compared to the coefficients of the corrected equation (column 1 or 2). It can be readily seen that there is a small downward bias for the local economy coefficients and a small upward bias for the Israeli economy coefficients. These directions of the bias are consistent with the afore-cited selection patterns. In terms of equation (23) the term $\frac{\partial c_i}{\partial x_k}$ is positive locally, negative in Israel.

This small coefficient bias, given by $\left[\frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}\right] \frac{\partial \lambda}{\partial c_i} \frac{\partial c_i}{\partial x_k}$ [see equation (23)], should not to be confused with the mean wage premium due to selection, given by $\frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}\lambda(c_i)$ [see equation (30)]. The estimates in Table 3 imply that the term $\sigma_{ii} - \sigma_{ij}$ is positive and substantial. The small bias in the estimated coefficients is due to the size of the term $\frac{\partial \lambda}{\partial c_i} \frac{\partial c_i}{\partial x_k}$. In contrast, the estimate of $\lambda(c_i)$ is such that the mean wage premium is sizeable.

Type of Self Selection. Tables 2 and 3 above report estimates of the unobserved skills variance-co-variance matrix ($\sum$). These allow for the analysis of the self-selection process on unobservables. As discussed in the preceding sub-section, a key issue is the relationship between the correlation of the unobserved skill distributions in the two locations (ρ_{ij}) and the relative skill standard deviations $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$. The relevant relation is given as follows:

$$\frac{\rho_i}{\rho_j} = \frac{\left(\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}} - \rho_{ij}\right)}{\left(\frac{1}{\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}} - \rho_{ij}\right)} \cdot \left(\frac{1}{\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}}\right) \tag{31}$$

The results of estimation indicate that:

(i) The correlation $\rho_{israel,local}$ is lower than the ratio of standard deviations $\frac{\sqrt{\sigma_{israel}}}{\sqrt{\sigma_{local}}}$.

(ii) The variance in local employment is higher than that of employment in Israel ($\sigma_{local} > \sigma_{israel}$).

Figure 1 illustrates the estimated relation. It depicts equation (31) in two panels. In panel (a)) the ratio $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ is held fixed at three different values. In panel (b) the moment ρ_{ij} is held fixed at four different values. The vertical axis measures the selection outcome $\frac{\rho_1}{\rho_j}$. The horizontal axis measures the relevant second moment of the unobserved distributions, $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ or ρ_{ij}.[25]

Figure 1

In panel (a) the purple line depicts the relation when $\frac{\sqrt{\sigma_{local}}}{\sqrt{\sigma_{Israel}}} = 0.8$, as estimated; points 1 and 2 on the line show the point estimates of the corresponding columns in Table 3. The black line shows the limit case of $\frac{\sqrt{\sigma_{local}}}{\sqrt{\sigma_{Israel}}} = 0.999$ and the blue line shows the limit case of $\frac{\sqrt{\sigma_{local}}}{\sqrt{\sigma_{Israel}}} = 0.001$.

The figure features the different cases as follows. The case of absolute advantage is given by the region where $\rho_{ij} < 0$ and there is positive selection in both locations. The comparative advantage case breaks down into two parts: left of the intersection of the lines with the ρ_{ij} axis, $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}} < \rho_{ij}$ and there is positive selection in both locations; right of that point $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}} > \rho_{ij}$ and selection is negative in location 1 ($\rho_1 < 0$), positive in the other location ($\rho_j > 0$), and so $\frac{\rho_i}{\rho_j} < 0$. When the dispersion ratio $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ is high (low) small changes in the estimates of $\frac{\rho_i}{\rho_j}$ make a big (small) difference in the implied correlation ρ_{ij}.

The actual estimates, depicted by the two points on the purple line, indicate that here the comparative advantage case obtains, with positive self-selection in both locations, implying positive correlation of the unobserved skill distributions (case (iii) in Section 6 above). As $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ is relatively high, there is a big sensitivity of the implied correlation ρ_{ij} to changes in the estimate of $\frac{\rho_i}{\rho_j}$. There is positive selection on unobservables of both movers and stayers, and the positive correlation between their unobserved skills is not too high as to overturn this result.

In panel (b) the purple line depicts the relation when $\rho_{local,Israel} = 0.7$, as estimated; points 1 and 2 on the line show the point estimates of the corresponding columns in Table 3. The yellow line shows the zero correlation case. The black line shows the limit case of $\rho_{local,Israel} = 0.999$ and the blue line shows the limit case of $\rho_{local,Israel} = -0.999$. When the correlation ρ_{ij} is very high, positive or negative, changes in the estimates of $\frac{\rho_i}{\rho_j}$ make big differences in the implied dispersion ratio $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$. When the correlation is zero

[25] Recall that the Heckman regression estimates produce an estimate of $\frac{\rho_1}{\rho_2}$ and of $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}$ from which ρ_{12} is deduced.

($\rho_{ij} = 0$) the relationship turns into the 45 degree line, i.e., $\frac{\rho_i}{\rho_j} = \frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$.

The results of estimation are once more depicted by the two points on the purple line, indicating the comparative advantage case with positive self-selection in both sectors, implying a dispersion ratio of the unobserved skill distributions that is lower than the correlation.

The unobserved skill distributions. The estimates of the parameters of the unobserved distributions are plotted in Figure 3. The figure shows the residuals from the Heckman log wage regressions reported in Table 3, column 1, together with fitted normal densities, for local and for Israel employment. The last panel of the figure shows the two fitted densities in one graph.[26]

Figure 2

Taking together the results depicted by Figures 1 and 2, one can see that they are reasonable: the positive correlation, which is not too high, is probably due to the fact that local and Israeli occupational tasks differed, as discussed above. Israeli tasks require skills that are less dispersed than those in the more high-skilled occupations of local employment – an "anybody can do it" effect – hence the lower variance in Israel employment.

Borjas, Kauppinen, and Poutvaara (2019) show that the distribution of unobservable skills for group i stochastically dominates that for group j when (using the notation here) $\rho_{ij}\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}} > 1$, i.e., $\rho_{ij} > \frac{\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}}}$, the comparative advantage case (i) of sub-section 6.1 above. The findings here are $\rho_{Israel,local} = 0.63$, $\sqrt{\sigma_{Israel}} = 0.34$, $\sqrt{\sigma_{local}} = 0.42$. Hence there is no stochastic dominance in unobservable skills. This is so as the variance of local unobserved skills is higher than that of the comparable Israeli skill distribution (and, moreover, the correlation of unobserved skills across locations is not sufficiently high).

7 What Do We Learn from the Natural Experiment?

I put together the lessons we learn from this natural experiment for the mover – stayer decision in the context of poor and rich economies differences.

7.1 A Graphical Representation

A graphical representation can clarify the moving decision and its various components. Consider the following regression equation implemented for

[26]Visual inspection of the last panel in the figure is not sufficient to determine self-selection patterns as these are determined by the relation between the ratio of standard deviations $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}$,which can be roughly seen, to the correlation $\rho_{1,2}$,which cannot be deduced from the figure. Figure 3 below facilitates the view of this relation.

the current case:[27]

$$\begin{aligned}\ln t_{Israel} &= \mu_{Israel} + \frac{\sigma_{local,Israel}}{\sigma_{local}}(\ln t_{local} - \mu_{local}) + \varepsilon_{Israel} \qquad (32)\\ &= \left(\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}\right) + \frac{\sigma_{local,Israel}}{\sigma_{local}}\ln t_{local} + \varepsilon_{Israel}\end{aligned}$$

where:

$$var\ \varepsilon_{Israel} = \sigma_{Israel}[1 - \frac{\sigma^2_{local,Israel}}{\sigma_{local}\sigma_{Israel}}]$$

In log tasks ($\ln t_j - \ln t_i$) space this regression is shown in the following figure (based on the discussion in Heckman and Sedlacek (1985, Figures 1 and 2)).

Figure 3

To understand the figure note the following elements:

(i) For any given worker, the log task value ($\ln t_{local}$) in the local location is given by a value on the horizontal axis.

(ii) The (red) regression line gives the linearly predicted log task value in the Israel location, i.e., predicted $\ln t_{Israel}$. It has the intercept given by $\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}$[28], and the slope given by $\frac{\sigma_{local,Israel}}{\sigma_{local}}$.

Actual values lie along the normal distribution around the regression line, as shown in two places in the figure; note that the distributions plotted relate to the vertical $\ln t_{Israel}$ values. The data points are distributed – conditional on the $\ln t_{local}$ value – with $var\ \varepsilon_{Israel}$.

The regression line and the normal distribution are plotted using the point estimates of the parameters and second moments reported in column 1 of Table 3.

(iii) The other line in the figure is the 45 degree line serving as the line of equal income ($\ln w_{local} = \ln w_{Israel}$).[29] It starts from a negative intercept as $\pi_{Israel} > \pi_{local.}$

This 45 degree line is key for the moving decision: when the worker has a value below this line he chooses the local economy; above it, he chooses to work in Israel. Hence, the fraction of workers choosing to move is the part of the normal distribution above the line, while the part below it is the fraction of stayers.

---

[27]Derived from multiplying both sides of the equation $\ln t_i = \mu_i + u_i$by $\frac{\sigma_{ij}}{\sigma_{ii}}$ and subtracting from $\ln t_j$.

[28]I use the point estimates of the coefficients (from Table 3, column 1), and the sample means of the X variables, to generate μ_{local} and μ_{Israel}. I adopt the normalization of $\beta_0 = 0$.

[29]Equal income means $\ln w_i = \ln w_j$ or $\ln \pi_i + \ln t_i = \ln \pi_j + \ln t_j$. Hence it is given by $\ln t_j = \ln \pi_i - \ln \pi_j + \ln t_i$.

Using the actual estimates from column 1 of Table 3, three major features of the analysis are manifested in the figure.

Country/moving premium. The Israeli economy, being more productive, has a higher task price i.e., $\pi_{Israel} > \pi_{local}$. Hence the (black) line of equal income starts from below 0.[30]

Negative selection on observables. The estimates indicate $\mu_{Israel} < \mu_{local}$ as the Israeli economy has low returns for skills (education and experience). Hence the intercept of the regression line[31] is negative. Thus, moving along the (red) regression line, the workers with relatively low $\ln t_{local}$ choose to work in Israel, as in that region the regression line lies above the 45 degree line; with relatively high $\ln t_{local}$ workers choose to work locally.

Positive selection on unobservables. The figure illustrates the positive selection on unobservables in each location.[32]Graphically $\frac{\sigma_{local,Israel}}{\sigma_{local}}$ is positive so the regression slope is positive. Note that when individuals are classified according to their task value, the fraction of people working locally increases as the local task level increases. In other words, as one moves up the $\ln t_{local}$ axis, the fraction of workers selecting the local economy rises. A similar graph with $\ln t_{israel}$ on the horizontal axis (not plotted here) would show a similar selection effect in the Israeli economy.

7.2 The Key Findings

The findings here are very much in the ballpark of what recent studies of other episodes of movers to rich economies have found in terms of magnitudes, and are consistent with recent findings in the development accounting literature. The contribution of the current analysis is twofold: first, it identifies the specific or "pure" roles of income differences in the move from a poor to a rich economy; second, it shows that the wage gains to movers are actually mitigated by the human capital differences flagged by the recent development accounting literature.

In disentangling the different effects at play, the key lessons are as follows:

(i) TFP and capital stock differences are indeed large; there are substantial productivity differences (the z_i) in favor of the rich economy, operating to raise the wages of movers.

(ii) The gains are offset to a large extent by big disparities in skill premia ($\boldsymbol{\beta}\mathbf{X}$), which reflect substantial human capital differences. The movers do

[30]The intercept is given by $\ln \pi_{local} - \ln \pi_{Israel}$.

[31]Given by $\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}$

[32]In terms of equation (30) this means that in each sector

$$E\left(\ln w_i \mid \{\ln w_i + \ln [1 - k_i(\mathbf{L})] > \ln w_j + \ln \left[1 - k_j(\mathbf{L})\right]\}\right) > E(\ln w_i).$$

not gain from the human capital differentials across countries, as they stay with their poor country skills.

(iii) While negative selection on observables plays a substantial role (as manifested in point ii), the positive selection on unobservables is not very important quantitatively.

7.3 Counter-Factuals

One question of interest is to consider how moving behavior would change following changes in the observed skill premia and in the unobserved skills distributions. The model is able to predict the size of moving when key parameters (π, μ), determining first moments, change. But changes in second moments $(\sigma_{ii}, \sigma_{ij})$ lead to ambiguous outcomes, as contradictory effects are at play. These results can be seen in the graphical framework of Figure 3 as follows.

Moving unambiguously rises when:

a. The moving premium rises, i.e., when $\frac{\pi_{host}}{\pi_{local}}$ rises. The line of equal income shifts downwards (i.e., the black line moves down). Fewer workers choose the local economy and more move.

b. When skill premia in the host economy (μ_{host}) rises or skill premia in the local economy (μ_{local}) fall. This raises the intercept, shifting the regression line upwards (the red line in the figure). More workers choose foreign employment.

The change in moving is ambiguous when the following changes in the unobserved skills distributions take place:

a. When the local (source economy) distribution becomes more dispersed, i.e., σ_{local} rises, the intercept rises and the slope declines so the regression line rises and flattens. In addition, the variance of the normal distribution around the line rises. The overall effect is ambiguous.

b. When the co-variance of the skills across the two economies declines, i.e., $\sigma_{local,host}$ falls, the same happens: the regression line shifts up and flattens and the normal distribution becomes more dispersed. Again, the overall effect is ambiguous.

c. When the host location distribution becomes less dispersed, i.e., σ_{host} falls, the variance of the normal distribution falls. The overall effect is once more ambiguous.

This analysis implies that government policy would generate unambiguous moving changes if it affects task prices, for example through taxation. Any policy which affects skills, such as education policy, has more complex outcomes. In particular, policy influencing $\sum$ has ambiguous moving outcomes.

8 Conclusions

The move from poor to rich countries is a prevalent and important phenomenon. This paper has used a natural experiment that facilitates the study of this move without confounding factors. The substantial productivity and human capital differences across economies turn out to play opposing roles, with higher productivity in the rich, host economy raising movers' wages, and the lower human capital component, embodied in the movers, operating to lower them. A challenge for future research is to undertake similar decompositions in the prevalent cases whereby confounding factors are present and to try to disentangle their relative effects.

References

[1] Acemoglu, Daron, and Melissa Dell. 2010. "Productivity Differences Between and Within Countries" **American Economic Journal: Macroeconomics**, 2(1): 169–88.

[2] Angrist, Joshua D. , 1995. "The Economic Returns to Schooling in the West Bank and Gaza Strip, " **American Economic Review** 85, 5, 1065-1087.

[3] Autor, David and Michael J. Handel, 2013. "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages," **Journal of Labor Economics,** 31, 2, S59-S96.

[4] Bartram, David V., 1998. "Foreign Workers in Israel: History and Theory," **International Migration Review** 32, 2, 303-325.

[5] Borjas, George J., Ilpo Kauppinen, and Panu Poutvaara, 2019. "Self-Selection of Emigrants: Theory and Evidence on Stochastic Dominance in Observable and Unobservable Characteristics," **Economic Journal** 129,617, 143–171.

[6] Caselli, Francesco and Antonio Ciccone, 2019. "The Human Capital Stock: A Generalized Approach: Comment." **American Economic Review** 109,3, 1155-74.

[7] Central Bureau of Statistics, 1996. **Labor Force Survey in Judea, Samaria and the Gaza Strip.**

[8] Clemens, Michael A., Claudio Montenegro, and Lant Pritchett, 2019. "The Place Premium: Bounding the Price Equivalent of Migration Barriers," **The Review of Economics and Statistics** 101, 2, 201- 213.

[9] Dao, Thu Hien, Frederic Docquier, Chris Parsons, and Giovanni Peri, 2018. " Migration and Development: Dissecting the Anatomy of the Mobility Transition, " **Journal of Development Economics 1**32, 88–101.

[10] Dustmann, Christian and Joseph-Simon Gorlach, 2016. "The Economics of Temporary Migrations" **Journal of Economic Literature** 54,1, 98-136.

[11] Dustmann, Christian and Costas Meghir, 2005. "Wages, Experience and Seniority," **Review of Economic Studies,** 72,1,77-108.

[12] Dustmann, Christian and Ian Preston, 2019. "Free Movement, Open Borders, and the Global Gains from Labor Mobility," CReAM, UCL discussion paper 04/19.

[13] Heckman, James J., 1979. "Sample Selection Bias as a Specification Error," **Econometrica** 47, 1, 153-161.

[14] Heckman, James J. and Guilherme L. Sedlacek, 1985. " Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," **Journal of Political Economy** 93,6, 1077-1125.

[15] Hendricks, Lutz and Todd Schoellman, 2018. "Human Capital and Development Accounting: New Evidence from Wage Gains at Migration,"**Quarterly Journal of Economics** 133, 2, 665–700.

[16] Jones, Benjamin F., 2014. "The Human Capital Stock: A Generalized Approach," **American Economic Review** 104, 11, 3752-77.

[17] Jones, Benjamin F., 2019. "The Human Capital Stock: A Generalized Approach: Reply," **American Economic Review**, 109, 3, 1175-95.

[18] Jones, Charles I., 2016. "The Facts of Economic Growth," in John B. Taylor and Harald Uhlig (eds.), **Handbook of Macroeconomics volume 2,** 3–69, Amsterdam, Elsevier.

[19] Kennan, John, 2013. "Open Borders," **Review of Economic Dynamics** 16, L1–L13.

[20] Kennan, John and James R. Walker, 2011. "The Effect of Expected Income on Individual Migration Decisions," **Econometrica** 79 (1), 211–251.

[21] Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, 2018a. "Life Cycle Wage Growth across Countries," **Journal of Political Economy** 126, 2, 797-849.

[22] Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, 2018b. "Life-Cycle Human Capital Accumulation across Countries: Lessons from US Immigrants," **Journal of Human Capital** 12, 2, 305-342.

[23] Razin, Assaf, and Sadka, Efraim, 1993. **The Economy of Modern Israel: Malaise and Promise.** University of Chicago Press, Chicago.

[24] Roy, Andrew D., 1951."Some Thoughts on the Distribution of Earnings," **Oxford Economic Papers** 3, 135-146.

[25] Semyonov, Moshe and Lewin-Epstein, Noah, 1987. **Hewers of Wood and Drawers of Water.** Ithaca, NY: Industrial and Labor Relations Press, 1987.

[26] Wooldridge, Jeffrey M., 2015. "Control Function Methods in Applied Econometrics," **The Journal of Human Resources** 50,2,420-445.

[27] Yashiv, Eran, 2000. "The Determinants of Equilibrium Unemployment," **American Economic Review** 90,1297-1322.

[28] Young, Alwyn, 2013. "Inequality, the Urban-Rural Gap, and Migration," **Quarterly Journal of Economics** 128, 4, 1727–1785.

9 Appendix

Derivation of GDP per Worker and Wages

I postulate a Cobb Douglas production function:

$$Y_i = K_i^{\alpha}(A_i T_i)^{1-\alpha} \tag{33}$$

This implies:

$$\frac{K_i}{Y_i} = \frac{K_i}{K_i^{\alpha}(A_i T_i)^{1-\alpha}} = \frac{K_i^{1-\alpha}}{(A_i T_i)^{1-\alpha}} \tag{34}$$

$$\left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} = \left(\frac{K_i}{A_i T_i}\right)^{\alpha} \tag{35}$$

GDP per worker is thus given by:

$$\begin{aligned} \frac{Y_i}{L_i} &= \left(\frac{K_i}{L_i}\right)^{\alpha} (A_i \frac{T_i}{L_i})^{1-\alpha} \\ &= \left(\frac{K_i}{A_i T_i}\right)^{\alpha} A_i \frac{T_i}{L_i} \end{aligned} \tag{36}$$

Using the relation $T_i = L_i \widetilde{T}_i$ and equation (35):

$$\frac{Y_i}{L_i} = \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} A_i \widetilde{T}_i \tag{37}$$

Define:

$$z_i \equiv \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} A_i \tag{38}$$

So:

$$\ln \frac{Y_i}{L_i} = \ln z_i + \ln \widetilde{T}_i \tag{39}$$

Assuming workers are paid their marginal products, real wages per worker in this set-up are given by:

$$\begin{aligned} \ln w_i &= \ln(1-\alpha) + \ln \frac{Y_i}{L_i} \\ &= \ln(1-\alpha) + \ln z_i + \ln \widetilde{T}_i \end{aligned} \tag{40}$$

10 Tables

Table 1
Sample Statistics
Palestinian Male Workers, 1987

variable	**Working in the Local Economy**	**Working in Israel**
N	7,206	11,670
wage (hourly, in logs)	-2.82 (0.44)	-2.73 (0.35)
education (in years)	8.9 (4.4)	7.8 (3.9)
experience (in years)	19.0 (13.1)	18.0 (13.1)
regions of residence		
Jenin	8 %	10%
Nablus	17%	6%
Tulkarm	7 %	14 %
Ramallah	17%	13%
Jordan valley	2 %	1 %
Bethlehem	11%	12%
Hebron	20%	17%
Rafiah	2%	4 %
Gaza	13%	15%
Khan Yunis	4%	9%
rural residence	41%	61 %
urban residence	47 %	22 %
refugee camp residence	12 %	17 %
married	68%	67 %

Notes:

1. For log wages, years of education and years of experience, the table reports mean of variables with standard deviations in parentheses.
2. The region of residence, type of residence and married numbers are percentage of workers out of total sample in the column.

Table 2: The Selection Equation
Probability of selection of employment in Israel

	(1)	(2)
constant	0.68	1.51
	(0.09)	(0.10)
education	−0.08	−0.08
	(0.003)	(0.003)
experience	−0.03	−0.04
	(0.003)	(0.004)
experience^2/100	0.02	0.04
	(0.005)	(0.005)
Refugee camp		−0.98
		(0.03)
Urban		−0.36
		(0.03)
Married		0.17
		(0.03)
Jenin	0.84*	0.18
Nablus	0.09	−0.33*
Tulkarm	1.15*	0.66*
Ramallah	0.55*	−0.08
Bethlehem	0.78*	0.26*
Hebron	0.56*	0.08
Rafiah	1.15*	0.95*
Gaza	0.82*	0.79*
Khan Yunis	1.31*	1.06*

Notes:

1. The equation relates to the probability of selection of employment in Israel. The specifications are elaborated in Section 4.4; see, in particular, equation (24).

2. The sample includes all wage earners except those with hourly wages below 0.1 NIS and above 11.5 NIS (cutting lowest 1% and highest 0.2 %).

3. The number of observations is 11,670.

4. Standard errors of the coefficients are in parentheses, except for the region of residence variables where a star denotes significance at 1%.

5. The equations included dummy variables for quarters, which are not reported.

6. The baseline region of residence is the Jordan valley.

Table 3
The Wage Equation
Dependent variable: log real hourly wage

exclusion restrictions	**(1) one**		**(2) three**		**(3) OLS**	
	Local	Israel	Local	Israel	Local	Israel
constant	-3.81	-3.10	-3.66	-3.09	-3.58	-3.09
	(0.04)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
education	0.043	0.011	0.038	0.012	0.036	0.013
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
experience	0.037	0.017	0.035	0.017	0.035	0.018
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
experience^2	-0.048	-0.027	-0.046	-0.027	-0.046	-0.028
(/100)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
ρ_i	0.35	0.13	0.14	0.04		
$\sqrt{\sigma_{ii}}$	0.42	0.34	0.41	0.34	0.40	0.34
$\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$	0.81		0.84			
$\frac{\rho_i}{\rho_j}$	0.37		0.29			
ρ_{ij}	0.63		0.71			
Wald test	2,009	1,457	2,127	1,392		
	(0.00)	(0.00)	(0.00)	(0.00)		
n	7,206	11,670	7,206	11,670	7,206	11,670

Notes:

1. The sample includes all wage earners except those with hourly wages below 0.1 NIS and above 11.5 NIS (cutting lowest 1% and highest 0.2%).

2. The specifications are discussed in Section 4.4; see in particular equation (25). Column 1 uses the geographical location varaible for the exclusion restriction. Column 2 adds the type of residence and marital status to the exclusion restrictions.

3. n is the number of observations in the regression.

4. Standard errors of the coefficients are in parentheses.

5. The regressions included dummy variables for quarters, which are not reported.

6. The Wald test is distributed χ^2. P-values appear in parentheses.

7. The second moment estimates use the relations:

$$\begin{aligned} \rho_i &= \left[\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}} - \rho_{ij} \right] \frac{\sqrt{\sigma_{jj}}}{\sigma^*} \\ \rho_j &= \left[\frac{\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}}} - \rho_{ij} \right] \frac{\sqrt{\sigma_{ii}}}{\sigma^*} \end{aligned}$$

Table 4
Decomposition of Mean Wages and of the Mean Wage Differential

a. Mean Log Wages

$$\begin{aligned}\overline{\ln w_{local}} &= \widehat{k}_{local} + \widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} + \widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} \\ \overline{\ln w_{Israel}} &= \widehat{k}_{Israel} + \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel} + \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}\end{aligned}$$

	local	**Israel**	**difference**
mean $\overline{\ln w}$ actual	-2.82	-2.73	-0.09
$\widehat{k}$	-3.806	-3.096	-0.71
$\widehat{\boldsymbol{\beta}}\overline{\mathbf{X}}$	0.965	0.364	0.60
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\overline{\lambda}}$	0.021	0.002	0.02

–

b. The Mean Wage Differential I

$$\begin{aligned}\overline{\ln w_{local}} - \overline{\ln w_{Israel}} = & \ \widehat{k}_{local} - \widehat{k}_{Israel} \\ & + \overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel}) \\ & + \widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel}) \\ & + \widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}\end{aligned}$$

$\overline{\ln w_{local}} - \overline{\ln w_{Israel}}$	-0.09
$\widehat{k}_{local} - \widehat{k}_{Israel}$	-0.71
$\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$	0.54
$\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$	0.06
$\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}$	0.02

c. The Mean Wage Differential II

$$\begin{aligned}
\overline{\ln w_{local}} - \overline{\ln w_{Israel}} &= \widehat{k}_{local} - \widehat{k}_{Israel} \\
&\quad -\{\overline{\mathbf{X}}_{local}(\widehat{\boldsymbol{\beta}}_{Israel} - \widehat{\boldsymbol{\beta}}_{local}) + \widehat{\boldsymbol{\beta}}_{Israel}(\overline{\mathbf{X}}_{Israel} - \overline{\mathbf{X}}_{local})\} \\
&\quad +\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}
\end{aligned}$$

$\overline{\ln w_{local}} - \overline{\ln w_{Israel}}$	-0.09
$\widehat{k}_{local} - \widehat{k}_{Israel}$	-0.71
$\overline{\mathbf{X}}_{local}(\widehat{\boldsymbol{\beta}}_{Israel} - \widehat{\boldsymbol{\beta}}_{local})$	-0.59
$\widehat{\boldsymbol{\beta}}_{Israel}(\overline{\mathbf{X}}_{Israel} - \overline{\mathbf{X}}_{local})$	-0.01
$\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}$	0.02

Table 5
Education and Age Distributions by Work Locations

a. Schooling Groups

years	0	1- 4	5-6	7-8	9-12	13+
Israel	7%	9%	22%	17%	38%	7%
Local	6%	9%	19%	13%	31%	22%

b. Age Groups

years	18-24	25-34	35-44	45-54	55-64
Israel	39%	33%	15%	8%	5%
Local	28%	33%	21%	12%	6%

Note:
Sample is the same as in Tables 1-3.

Table 6
Industry and Occupation Distributions by Work Locations

a. Industry Distributions

industry	**Local**	**Israel**
agriculture	3%	11%
manufacturing	24%	20%
construction	22%	50%
commerce	5%	9%
government	34%	6%
transportation	6%	2%
personal services	3%	3%
finance	1%	0%

b. Occupation Distributions

occupation	**Local**	**Israel**
academic	7%	0%
professionals	13%	1%
managers	1%	1%
clerical workers	9%	1%
agents, sales and service	3%	2%
skilled jobs in agriculture	8%	12%
manufacturing and construction skilled jobs	37%	42%
unskilled	22%	42%

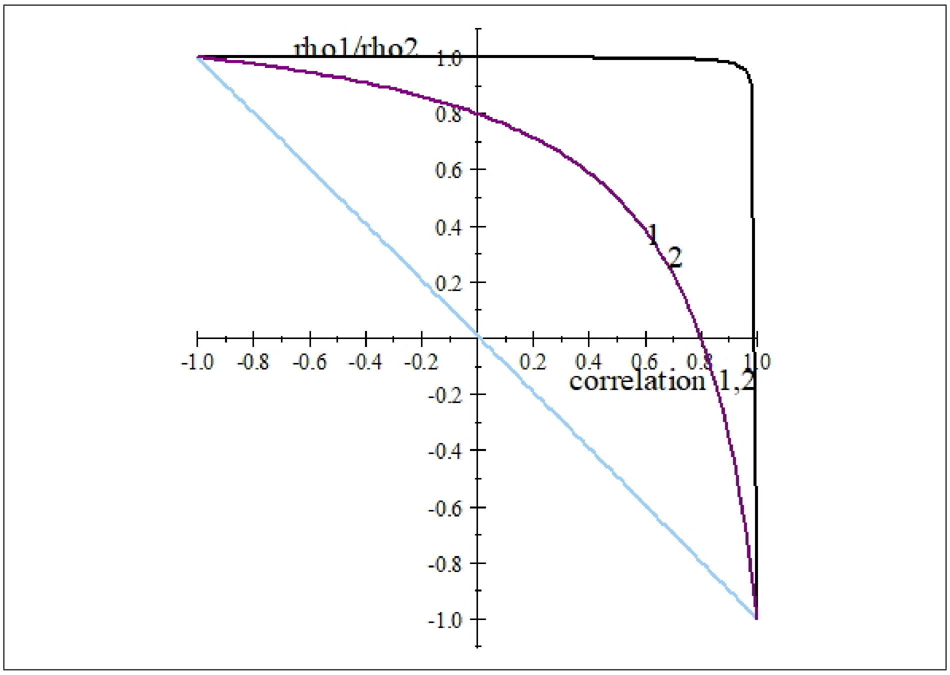
Note:
Sample is the same as in Tables 1- 3.

11 Figures

Figure 1
Patterns of Self Selection on Unobservable Skills

A. The Relation Between $\frac{\rho_1}{\rho_2}$ and ρ_{12}

$$\frac{\rho_1}{\rho_2} = \frac{\left(\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} - \rho_{12}\right)}{\left(\frac{1}{\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}} - \rho_{12}\right)} \cdot \left(\frac{1}{\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}}\right)$$



Notes:

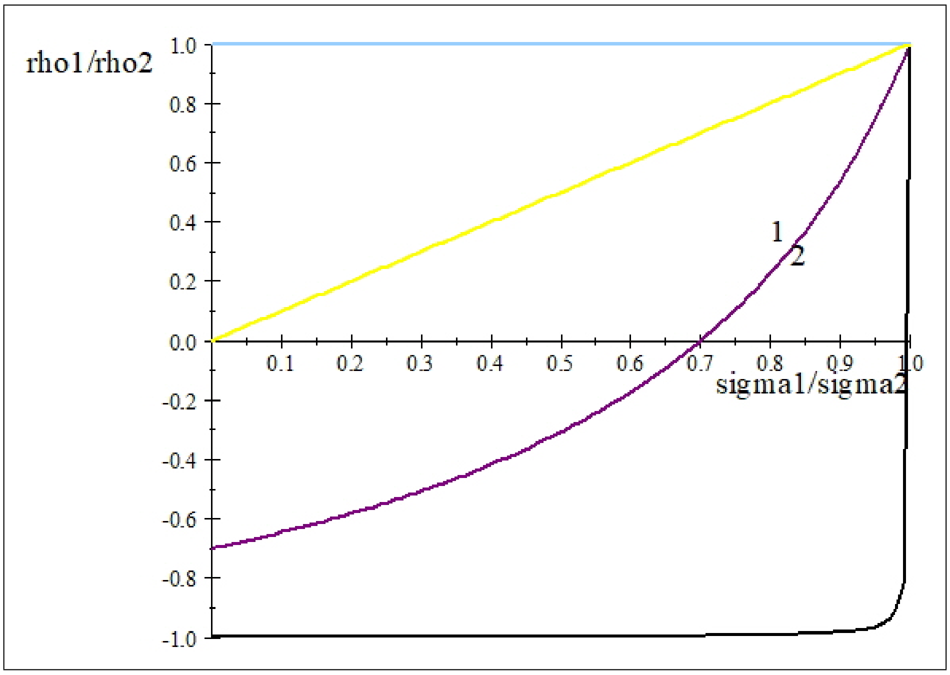
1. The vertical axis depicts $\frac{\rho_1}{\rho_2}$ where $\rho_i = correl(u_i, u_i - u_j)$, $i \neq j; i, j = 1, 2$; the horizontal axis depicts ρ_{12}.

2. The purple line depicts the relation when $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} = 0.8$ as estimated; the black line shows the limit case of $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} = 0.999$; the blue line shows the limit case of $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} = 0.001$.

3. Points 1 and 2 on the purple line show the point estimates of the corresponding columns in Table 3.

B. The Relation Between $\frac{\rho_1}{\rho_2}$ and $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}$

$$\frac{\rho_1}{\rho_2} = \frac{\left(\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}} - \rho_{12}\right)}{\left(\frac{1}{\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}} - \rho_{12}\right)} \cdot \left(\frac{1}{\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}}\right)$$



Notes:

1. The vertical axis depicts $\frac{\rho_1}{\rho_2}$ where $\rho_i = correl(u_i, u_i - u_j),\ i \neq j; i, j = 1, 2$; the horizontal axis depicts $\frac{\sqrt{\sigma_{11}}}{\sqrt{\sigma_{22}}}$.

2. The purple line depicts the relation when $\rho_{12} = 0.7$ as estimated; the black line shows the limit case of $\rho_{12} = 0.999$; the blue line shows the limit case of $\rho_{12} = -0.999$; the yellow line shows the case of $\rho_{12} = 0$.

3. Points 1 and 2 on the purple line show the point estimates of the corresponding columns in Table 3.

Figure 2
Unobserved Skills: Log Wages Residual Graphs

a. Workers in the Local Economy

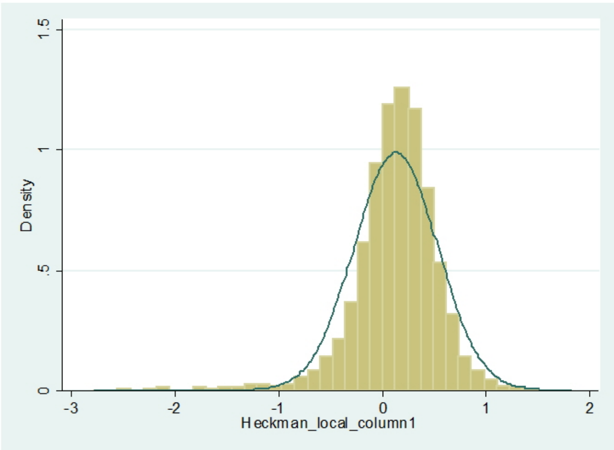


Table 3, col. 1, local, log wage residuals

b. Workers in the Israeli Economy

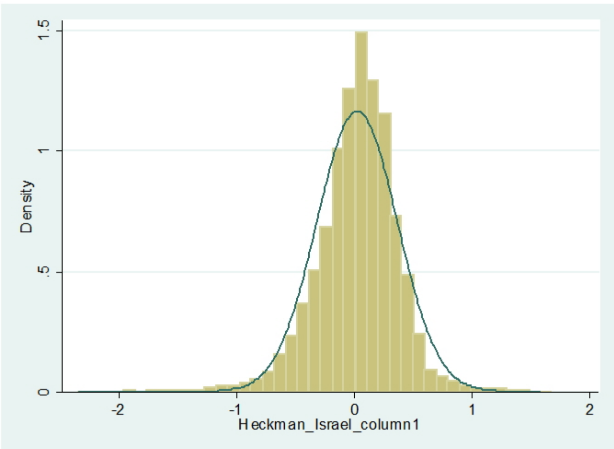
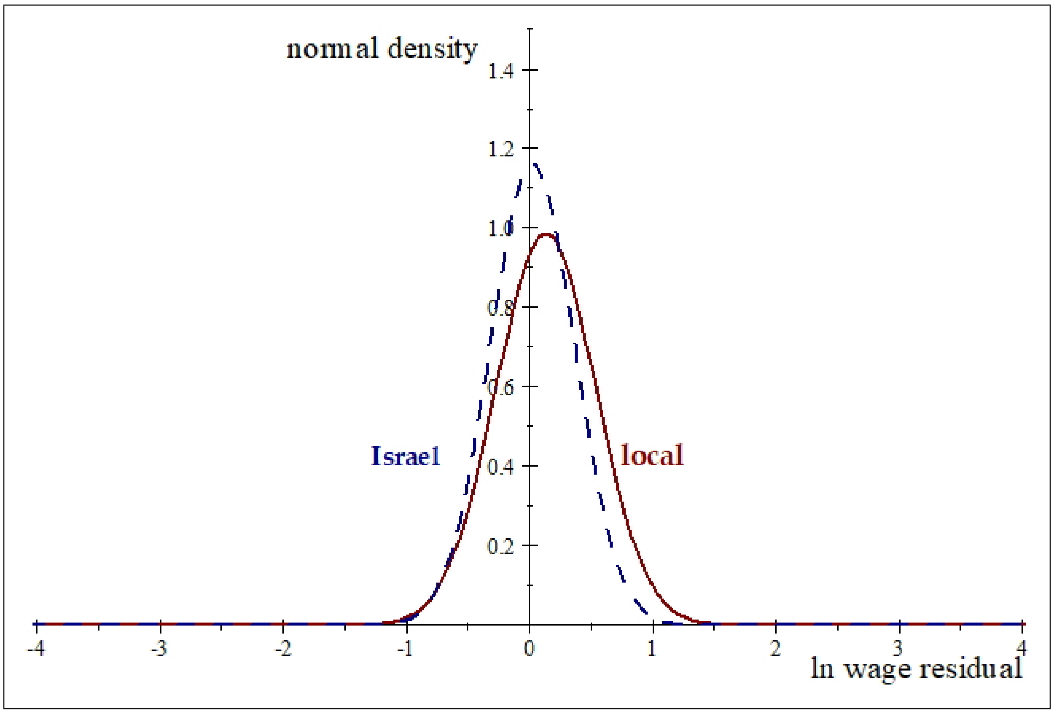


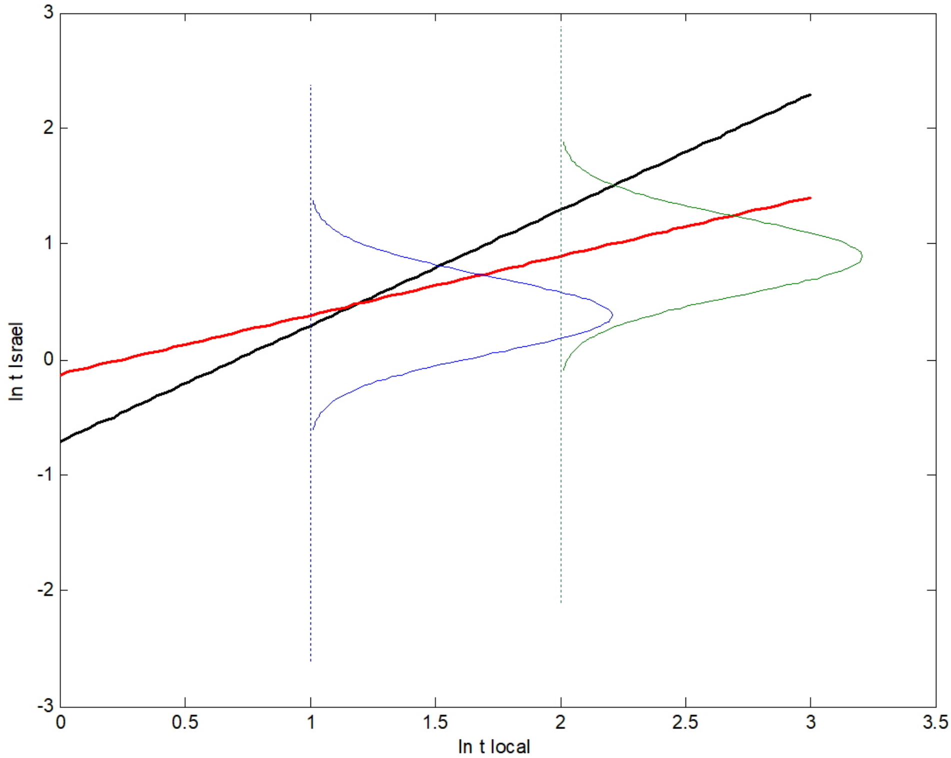
Table 3, col. 1, Israel, log wage residuals

c. Fitted Normal Densities of Residual Log Wages



Notes:
1. Local – red, solid; Israel – blue, dashed
2. The horizontal scale ranges in panel a and b are somewhat different.

Figure 3: The Move-Stay Decision



Notes:

1. Equation (32) is given by the red regression line, which is upward sloping. The intercept is given by $\left(\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}\right)$; the slope is given by $\frac{\sigma_{local,Israel}}{\sigma_{local}}$; values along the line are distributed with $var\ \varepsilon_{Israel}$.

2. The equal income line, $\ln w_{Israel} = \ln w_{local}$ is given by the black line. The intercept is given by $\ln \pi_{local} - \ln \pi_{Israel}$ and the slope is 1 (45 degree line).

3. Workers choose work in Israel when above the black line and work locally when below the black line.

4. The regression line and the normal distribution are plotted using the point estimates of the parameters and second moments reported in column 1 of Table 3.