Moving from a Poor Economy to a Rich One: The Roles of Incomes and Job Tasks *

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

January 12, 2020

Abstract

When a worker moves from a poor to a rich economy, what is gained by the move? The paper addresses this question using a self-selection model that caters for two empirically-important sets of feaures. First, it encomapsses notable facts concerning rich and poor countries income differences, recently characterized by the development accounting literature. Second, the model explicitly recognizes that movers and stayers face different job tasks requirements and different rewards for their skills in performing these tasks.

The paper makes use of a unique data set on Palestinian workers, working locally and in Israel, that allows to isolate the pure effects of income differences with no confounding factors, while encompassing the constraints placed on movers in terms of the human capital skills required.

The findings show that income differences affecting the choice to move to a rich economy are made up of elements, which operate in opposition. Productivity differences in favor of the richer economy, due to differences in TFP and in physical capital, are sizeable and operate to raise wages for movers. Lower returns to human capital and lower stocks of human capital, however, operate to lower wages for movers. This is due to negative selection on observables, with movers being offered low-skill tasks in the rich economy. The latter effect offsets, to a large or full extent, the former gain.

Key Words: movers, stayers, rich and poor economies, development accounting, pure effects, income differentials, task approach, skill bundle, Roy model, TFP differentials, human capital differences, selection.

JEL Codes: E24, F66, J24, J31, O15.

* *I am grateful to Francesco Caselli, Chad Jones, John Kennan, and David Lagakos for useful conversations, to Arnaud Maurel for econometric advice, to seminar participants in numerous places for feedback and suggestions on previous versions, and to Elad Demalach, Oriel Nofekh, and Alon Rieger for research assistance. Any errors are my own.*

† yashiv@tauex.tau.ac.il

**Moving from a Poor Economy to a Rich One:
The Roles of Incomes and Job Tasks**

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). A related and important complication is the fact that movers and stayers are typically constrained in terms of the jobs offered and the skills required.

This paper studies a unique case that allows to isolate the pure effects of income differences and that caters for the constraints placed on movers in terms of the human capital skills required. 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 were such that the other cited factors played 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 daily or weekly commute. Without the confounding factors, the decision to work in the richer economy can thus be estimated without bias.

I use a self-selection model catering for two empirically-important sets of features. First, it encompasses notable facts concerning rich and poor countries income differences, as characterized by recent papers in the development accounting literature. The latter suggests sizeable rich-poor countries income differences exist, while debating the relative weights of their various constituents. Second, it explicitly recognizes that workers face job tasks requirements and particular rewards for their skills in performing these tasks.

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

Specifically, the empirical work breaks down the wage differentials motivating the movers from poor to rich economies. It estimates and quantifies (i) the productivity advantage of the richer, host economy relative to the poorer, home economy; (ii) the differential returns to human capital across the two economies; (iii) the differences in the stocks of human capital across these economies; (iv) self-selection on observables and on unobservables.

I take the model to the data using repeated cross sections of a labor force survey, which sampled both movers and stayers within a unified setting. The data feature a high proportion of movers from the home population. I use two alternative estimation methodologies to examine wage regressions of movers and stayers. My findings offer a nuanced view of the gains to movers, as the pure effects of income differences in the choice to move to a rich economy are made up of diverse elements, operating in opposition. Productivity differences in favor of the richer economy, due to differences in TFP and in the stock and quality of physical capital, are sizeable and operate to raise wages. However, lower returns to human capital and lower stocks of human capital for movers, operate to lower wages. The latter is due to negative selection on observables of movers, who are being offered low-skill tasks in the rich economy. The latter effect offsets to large extent the former gain, sometimes erasing it. Self-selection on unobservables turns out to play a far smaller quantitative role. These findings are consistent with the recent development accounting literature in terms of the pattern of income differences across countries, and reveal large *gross* differences. But they do not confirm the claim that *net* gains of such a move are large, due to the afore-mentioned offset.

While the literature often looks at migration without disentangling the income differences motive from a plethora of other motives and suffers from potential misspecification and bias, the empirical work here does not suffer such bias. The paper shows that the task approach to labor market and human capital analysis is key in understanding the consequences of the poor to rich economy move. Tasks are tied to locations, and so workers choose a location-wage-task 'pack' that determines rewards to the skills bundled in the task. This constrains the human capital returns for movers.

In a review of migration, productivity and the labor market, Peri (2016) emphasizes, the importance of recognizing the role of tasks performed by migrants, especially manual tasks. The latter feature is particularly important for the non-college educated. He discusses the fact that employment in manual, low skill occupations is a salient feature among them, as it is in the case of Palestinian men discussed here. Dustmann and Frattini (2013) offer a detailed review of the relevant data for Europe. Dustmann, Frattini, and Preston (2013) find, using UK data, that migrants work, at least initially, in jobs and occupations which do not fit their observed skills, i.e., experience "downgrading." The analysis here can thus be placed in a broader context. It may pertain to many cases of ethnic minorities in advanced economies.

Workers belonging to such minorities commute to work in a rich economy and are demanded to perform low-skill tasks, as is the case here.

The paper proceeds as follows. Section 2 offers the background and context. Section 3 presents the model. Section 4 discusses the data and summary statistics. Section 5 presents the econometric methodologies and the results. Section 6 analyzes the components of movers-stayers wage differentials and their significance. Section 7 examines the skills and wage distributions and their connections to the moving decision. Section 8 interprets the results in terms of the development accounting literature. Section 9 concludes.

2 Background and Context

This paper relates to two strands of literature. The first is the development accounting literature, discussed in sub-section 2.1, which studies cross-country income differences. The second is the task-based approach to the labor market, discussed in sub-section 2.2, which emphasizes the analysis of employment, occupation, and wages from the viewpoint of worker tasks and the skills to undertake them.

2.1 Development Accounting

A key question in the development accounting (DA) literature is the relative importance of TFP versus human capital in accounting for cross-country income differences.

Caselli (2005) and Jones (2016) offer reviews of the evidence, documenting very substantial differences in GDP per worker across countries. Focusing on TFP differences, Jones (2016) offers a number of explanations, mostly having to do with misallocation. In particular, misallocation at the micro level shows up as a reduction in total factor productivity at the aggregate level. Banerjee and Moll (2010) offer explanations for the persistence of such misallocation. Acemoglu and Dell (2010) 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.

This literature reports a wide range of estimates for TFP and human capital shares, ranging from 20% to 80% of cross-country income differences for the latter, with TFP accounting for most of the complementary share.

Hendricks and Schoellman (2018, 2019), henceforth HS, make key contributions to the debate on the relative size of TFP vs human capital shares in accounting for cross-country differences. Their work presents evidence from the experiences of immigrants to the United States. The underlying logic is that "immigrants provide information because they enter the United States with the human capital they acquired in their birth country, but not their birth country's physical capital or TFP. Hence, their labor market performance in the United States conveys information about their human capital separated from the other two country-specific factors. On the other hand, working with immigrants presents two well-known challenges. First, immigrants are selected: their human capital is not the same as the human capital of a randomly chosen person in their birth country. Second, their labor market performance may not accurately reflect their human capital if skills transfer imperfectly across countries." (HS (2018, p.666)). Examining data on migration to the U.S., mostly from poor economies, they attribute around 60% to human capital differences and the remainder to TFP and physical capital-related differences.

HS (2019) deepen their afore-cited inquiry by catering for various features of the data: imperfectly substitutable skills (examining alternative values for the elasticity of substitution between skilled and unskilled labor); cross-country variation in the skill bias of technology; alternative sources of skill-biased technology variation; and alternative definitions of skilled and unskilled labor. They find that human capital accounts for between 50% and 75% of cross-country income gaps, in line with their afore-cited earlier findings. The share of output gaps due to TFP differences ranges between 36% and 42% across skill definitions, and the remaining 4% are attributed to physical capital (see Table 7 in HS (2019)).

A related issue in this literature pertains to the determinants of worker efficiency in the case of workers with different skills and imperfect substitution. The question is to what extent does worker efficiency reflect human capital characteristics of the workers themselves (such as education, training, traits, etc.) or the technology and institutions in their environment (such as the production technology chosen). See, for example, the debate and discussions in Ciccone and Caselli (2019) and Jones (2019).

There are papers in the migration literature, focusing on migration from poor to rich economies, which relate to similar questions. Dustmann and Preston (2019) offer a comprehensive review.

In a prominent contribution, Kennan (2013) presents a general equilibrium model, which is subsequently 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%.

In this paper I relate my findings to the questions of movers' wage gains, and their distinctive determinants in terms of human capital and TFP or physical capital. I suggest a mechanism, to account for the results, which has not been evaluated by the afore-going strands of literature.

2.2 Tasks and Skills

Acemoglu and Autor (2011) survey the task approach to labor market analysis. The background for this approach is the recognition that the standard Becker-Mincer model is not informative about the demand side of the labor market related to human capital. Thus, it does not model the factors that determine the skills that firms demand and how skill requirements change over time. The task approach literature analyzes job skill requirements. It classifies jobs according to their task requirements and considers the skills required to carry out these tasks. This approach offers a foundation for linking the aggregate demand for skills in the labor market to the skill demands of given jobs. It has been used to explore the links between technological change, changes in task inputs, and shifts in the wage structure.

Within this approach, Autor and Handel (2013) depart from the premise that, unlike investment such as education, job tasks are not fixed worker attributes, as workers can modify their task inputs by self-selecting into particular jobs. They use the Roy (1951) self-selection framework to analyze the relationship between tasks and wages. They note that their approach is motivated by the fact that "while workers can hold multiple jobs, they can supply tasks to only one job at a time. The indivisible bundling of tasks within jobs means that the productivity of particular task inputs will not necessarily be equated across jobs—and so the "law of one price" will not generally apply to the market rewards to job tasks." (page S90). Using U.S. job and task data, they test the model's predictions for this relationship, finding empirical support for the model.

In the model here I use the Roy (1951) framework as further developed by Heckman and Sedlacek (1985). As I discuss below, the task approach is crucial in understanding the empirical results, in particular in light of the afore-cited DA literature analysis. The review by Peri (2016), cited above, indicates that tasks may be relevant in many migration contexts. The idea that job tasks may be location-specific features in the rich, general equilibrium model of migration, empirically studied by Bryan and Morten (2019).

3 The Model

Given the afore-going discussion, the model needs to cater for the following features. Income differences between the two economies should play a major role; there should be a distinction between TFP and physical capital determinants and human capital determinants in forming these income differences; it needs to model the job tasks involved; and it needs to cater for self-selection. A suitable model is the Roy (1951) model, as further developed and implemented by Heckman and Sedlacek (1985). As is well known, this model has been applied to labor market issues on many occasions.

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. When coming to implement the model empirically, I use both the self-selection methodology proposed by Heckman (1979) as well as the more recent semi-parametric methodology of D'Haultfoeuille, Maurel, and Zhang (2018).

3.1 The Movers Decision

Tasks and production. There are two localities, indexed i, j, the richer, Israeli economy, and the poorer, Palestinian, local economy, in which workers can work. 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:

$$\ln t_i(\mathbf{S}) = \beta_{i,0} + \sum_h \beta_{h,i} S_h + u_i \tag{6}$$

where h is an index of skills.

Hence:

$$\begin{aligned} \ln w_i(\mathbf{S}) &= \ln \pi_i + \ln t_i(\mathbf{S}) \\ &= \ln \pi_i + \beta_{i,0} + \sum_h \beta_{h,i} S_h + u_i \end{aligned} \tag{7}$$

Travel and psychic costs. The individual worker has travel costs to work. These depend on a vector of variables related to location, to be denoted $\mathbf{L}$, and are formulated as a fraction $k_i(\mathbf{L})$ of wages. This corresponds to the situation whereby part of the worker's wage was used to pay for the work commute.

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

I discuss the $\mathbf{L}$ variables in the empirical work below.

In addition it is possible to think of psychic costs entailed in working in Israel, given the hostility between Israelis and Palestinians. This will be formalized as a multiplicative fixed cost, $\exp(\ln(1-\gamma_i))$,where $\gamma_i = \gamma$ in Israel and $\gamma_i = 0$ in the local economy.

Income maximization. An income-maximizing individual chooses location i if:

$$(w_i(1-k_i(\mathbf{L}))) \cdot \exp(\ln(1-\gamma_i)) > w_j(1-k_j(\mathbf{L})) \cdot \exp(\ln(1-\gamma_j)) \tag{9}$$

This can also be written as:

$$[\pi_i t_i(S)]\,[1-k_i(\mathbf{L})]\cdot \exp(\ln(1-\gamma_i)) > \left[\pi_j t_j(S)\right]\,\left[1-k_j(\mathbf{L})\right]\cdot \exp(\ln(1-\gamma_j)) \tag{10}$$

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 face two wages:

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

With these functional specifications, the following holds true:[3]

$$pr(i) = P\left(\ln w_i + \ln\left[1-k_i(\mathbf{L})\right] + \ln(1-\gamma_i) > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right] + \ln(1-\gamma_j)\right) = \Phi(c_i) \tag{12}$$

where

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

$\Phi(\cdot)$ the cdf of a standard normal variable. The proportion of workers in location i will increase as the relative task price $\ln\frac{\pi_i}{\pi_j}$ rises, as relative costs decline, i.e. as $\ln(1-\gamma_i) - \ln(1-\gamma_j) + \ln\frac{[1-k_i(\mathbf{L})]}{\left[1-k_j(\mathbf{L})\right]}$ rises, or as the relative mean task $\mu_i - \mu_j$ 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, however, 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

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

to the location – Israel or the local economy. Hence wage gains are going to be empirically examined across locations and pertain to Palestinian workers only, i.e., movers and stayers, not across workers of the different economies, Israelis and Palestinians (the object of study of the DA literature).

As a parametric specification of equation (2), assume a Cobb Douglas production function, with physical capital K, human capital T, and technology A to produce product output in location i:

$$Y_i = K_i^{\alpha}(A_i T_i)^{1-\alpha} \tag{13}$$

Define:

$$z_i \equiv \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} A_i \tag{14}$$

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

$$\ln \frac{Y_i}{L_i} = \ln z_i + \ln t_i \tag{15}$$

Note that Y_i should not be confused with GDP of the country. Hence Y_i can be, for example, the output in the agriculture and construction sectors in Israel, with the associated job tasks (t_i), not Israeli GDP.

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

$$\begin{aligned} \ln w_i &= \ln(1-\alpha) + \ln \frac{Y_i}{L_i} \qquad (16) \\ &= \ln(1-\alpha) + \ln z_i + \ln t_i \end{aligned}$$

Using equation (7) this means:

$$\ln \pi_i = \ln(1-\alpha) + \ln z_i \tag{17}$$

Workers can gain by a move to a richer economy with a higher level of z_i (and therefore also labor productivity $\frac{Y_i}{L_i}$). The worker gains because of work in an economy with higher levels of K and/or A, as seen in equation (14). In terms of the preceding analysis, this means that the richer economy has a higher level of π_i (see equation (17)). These, however, are not the only consequences for wages. Equation (7) has shown that the term $\sum_h \beta_{h,i} S_h + u_i$ will be important for wages too. This term expresses task performance through the bundle of skills (S_h) and the rewards to these skills $(\beta_{h,i})$.

4 The Data and Summary Statistics

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.[4] 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.[5] Hence, a key employment decision of the Palestinian male worker was the choice of employment location – Israel or the local economy. Men constituted 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 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.[6]

In this paper I use data on Palestinian workers in the period 1981-1987. In these years 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 commuted daily, though some stayed a few days a week in Israel.[7]

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 Israeli Central Bureau of Statistics (CBS), in local cur-

[4] Arnon, Luski, Spivak, and Weinblatt (1997, in particular Chapter 3) discuss the dependence of the Palestinian economy on Israel, especially in terms of the labor market. The following discussion is based on the above reference, and on Bartram (1998). For further details on the Palestinian labor market, see these references and Angrist (1995); for an analysis of the Israeli labor market, see Yashiv (2000).

[5] Table 1 below gives the numbers for the sample years 1981-1987.

[6] For details on developments over time in the Palestinian labor market, see the aforecited references.

[7] Semyonov and Lewin-Epstein (1987, pp.13-15) describe the organizational arrangements.

rency and current prices.[8]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 reports the average ratio was 13% in the 25 year period from 1994 to 2018.[9]

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).[10]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 included 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.

I use observations on Palestinian men[11] aged 18-64 from repeated cross sections of the TLFS in the years 1981-1987. This sample period precedes the uprising and the ensuing turbulence, described in sub-section **??** above.

Table 1 presents full sample statistics.

Table 1

The table shows that, for most, but not all, years, local workers (stayers) earned slightly lower wages[12] 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.

5 Methodology and Results

I estimate selection and wage equations for Palestinian men working in Israel and East Jerusalem as one location and working locally in the West

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

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

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

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

[12]Those wage differences are analyzed at length below.

Bank and Gaza as the other location. In what follows I discuss the uniqueness of this data set (5.1), the econometric methodologies (5.2), and the results (5.3).

5.1 The Pure Effects of Income Differences

It has, of course, been previously recognized that not taking into account factors that affect the moving decision and expected earnings can represent an important source of bias. The set up of the current paper precludes this possibility. In what follows I delineate such factors, which have been discussed in the literature; see, for a recent example, the analysis in Dao, Docquier, Parsons, and Peri (2018). Typically these factors drive the moving decision but are not taken into account, often because of lack of relevant data. As explained below, in the current case they do not play a role and hence their omission is not problematic.

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 in Israel, were paid for by the employers, partly or fully 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 may have 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 consecutive days at most. Hence this determinant had much less power, if at all.

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

$$M = f(w_i - w_j, \mathbf{X}) + \varsigma \tag{18}$$

where M is the moving decision, $w_i - w_j$ are wage differences, $\mathbf{X}$ is a vector of determinants such as the ones discussed above, and ς 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 often estimated, omitting at least some of the elements of $\mathbf{X}$.

5.2 Econometric Methodology

I use two alternative methods to estimate equations (11), for workers employed locally and those employed in Israel. These methods are elaborated in Appendix B; the following is a short summary.

5.2.1 Heckman Selection Method

The Heckman (1979) selection methodology is applied. 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.[13]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. In an alternative case, I use the above two variables, additionally, 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:

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

$$k_i(\mathbf{L}) = \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^i \tag{19}$$

where l is the region of the worker's residence, p is an index of regions, and so the l_p variables are the dummy variables for geographical regions or localities and θ_p is a coefficient to be estimated; the Y_n variables are type of residence and marital status, the 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 (12). Summary statistics of these variables appear in Table 1below.

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[14]. I also use indicator variables for the quarters.

The dependent variable in the wage equation is the log of hourly wages ($\ln w_i$), defined as the monthly wage divided by hours worked. The use of hourly wages is designed to avoid confounding the choice of work place with the choice of work time (hours or days).[15] Education (*educ*) and experience (*exp*) are defined in years. The benchmark specification reported below (using the estimates of column 1 in Table 2) features only the geographical exclusion restrictions.. The alternative specification, column 2 in Table 2 below, also includes the variables discussed above and contained in $\mathbf{L}$, so there are three exclusion restrictions. The specification of column 3 in Table 2 uses OLS to test for the effect of selection correction (running only the wage equation).

5.2.2 Semi-Parametric Estimation

I use the semi-parametric methodology proposed by D'Haultfoeuille, Maurel, and Zhang (2018) and D'Haultfoeuille, Maurel, Qiu, and Zhang, (2019) to estimate the model equations (11) without relying on exclusion restrictions. The background to this methodology is the finding that identification without instruments is possible. The key condition for that is that selection be independent of the covariates at infinity, i.e., when the outcome takes arbitrarily large values. If selection is indeed endogenous, one can expect the effect of the outcome on selection to dominate those of the covariates, for sufficiently large values of the outcome. This idea is implemented by using an estimator based on an extremal quantile regression, i.e., a quan-

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

[15]The sample does not include the lowest 1% and the 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).

tile regression applied to the upper tail of the outcome variable. Appendix B provides a more formal elaboration.

5.3 Results

Table 2 reports the full results of the Heckman methodology[16] using the two alternative specifications for the exclusion restrictions, and of OLS, for the year 1987, which has the highest data quality.

Table 2

The OLS estimates are relatively close to the Heckman selection-corrected ones, except for the intercept in Israel employment. The emerging picture across columns 1 and 2 is the same, but column 1 has higher point estimates in absolute value for the returns to skills. Overall, the differences in point estimates across specifications are not substantial.

Table 3 and Figure 1 present the results for the seven repeated cross-sections in the years 1981 to 1987, using this methodology with the exclusion restrictions set 1 of Table 2.

Table 3 and Figure 1

The main results to note from Tables 2 and 3 and Figure 1 are as follows.

(i) The constant of the equation, essentially capturing $z_i \equiv \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} A_i$, is much higher in Israel relative to the local economy.

(ii) The returns to education and experience are much lower in Israel than in the local economy.

(iii) The selection of work in Israel is negatively related to education, experience, refugee camp and urban residence, and is positively related to being married. The magnitudes of the region coefficients are reasonable; areas that are relatively more distant from Israeli employment locations have lower coefficients of Israel selection than regions, which are relatively closer.

Table 4 reports the results of the semi-parametric methodology discussed in sub-section 5.2.2 above. It presents the skill premia estimates, and repeats the results of the Heckman specification (of Table 3 above), for all years 1981-1987.

Table 4

[16] I include estimates of 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).

The table shows that, overall, the finding in point (ii) above holds true across all years and across the two estimation methodologies. This means that the returns to education and experience are found to be much lower in Israel than in the local economy. The semi-parametric estimates of returns to education and to experience in the local (Israeli) economy are somewhat lower (higher) than the Heckman estimates, hence the semi-parametric methodology points to a somewhat lower gap of the skill premia between the two economies.

I turn now to examine the implications of these results.

6 Components of Wage Differentials and Their Significance

Understanding the move to a rich economy, which is based solely on the wage differential between movers and stayers, requires analysis of its components.[17]

6.1 Decomposition of the Wage Differential

In Table 5 and Figure 2, 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 reported in Table 3.

Table 5 and Figure 2

The table and figure report the constituents of mean wages in each of the locations, using the following equations:

$$\begin{aligned} \overline{\ln w_{local}} &= \widehat{k}_{local} + \widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} + \widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} \\ \overline{\ln w_{Israel}} &= \widehat{k}_{Israel} + \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel} + \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}} \end{aligned} \tag{20}$$

where $\overline{\ln w_i}$ is the mean log hourly wage in economy i, $\widehat{k}_i = \ln \widehat{\pi}_i + \widehat{\beta_{i,0}}$ 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. The table and figure pertain to the period 1981-1987, using the Heckman methodology.

[17] Note that the wage differential analysis undertaken here pertains to Palestinian workers movers and stayers, not to native workers of the two economies.

Subsequently, Table 5 shows the mean wage differential between Palestinian workers in the Israeli economy and in the local economy ($\overline{\ln w_{local}} - \overline{\ln w_{Israel}}$), broken down into components, using the following equation.

$$\begin{aligned}\overline{\ln w_{local}} - \overline{\ln w_{Israel}} &= \widehat{k}_{local} - \widehat{k}_{Israel} \\ &\quad + \overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel}) + \widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel}) \\ &\quad + \widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}\end{aligned} \tag{21}$$

The components include the part due to differences in task prices plus the intercept of the task function ($\widehat{k}_{local} - \widehat{k}_{Israel}$); a part due to differences in skill premia across the two locations (($\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})\overline{\mathbf{X}}_{Israel}$); a part due to differences in skill levels across the two locations ($\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$); and a part due to differences in selection effects ($\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}$).

The key findings from the table and the figure are as follows.

The mean wage differential in the data. The data shows that the mean wage differential for Palestinian workers across locations $\overline{\ln w_{local}} - \overline{\ln w_{Israel}}$ is small and changes sign across years. It ranges between -0.08 and $+0.17$ log points.

Moving premium. The wage equation's intercept – reflecting the task price π_i and the task function intercept $\beta_{i,0}$ – is substantially higher in Israel. The $\widehat{k}_{local} - \widehat{k}_{Israel}$ difference ranges between -0.48 and -1.09 log points across the seven years of repeated cross sections. Note that this difference in baseline wages, or 'moving premium,' is much higher than the afore-cited difference in mean wages between Israel and local employment. Hence there is a large offset to the moving premium to which I turn now.

Skill premia.

The local returns to education and experience[18]are higher in the local economy, as seen in Tables 3 and 4 and in Figure 2.[19]Hence one gets $\widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel} >> 0$. This difference ranges between 0.60 and 1 log points across the sample years.

Equation (21) breaks this latter expression down into two components: the skill premia difference component $\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$ plays the major part, ranging between 0.54 and 0.88 across the sample years; the skill stocks component $\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$ ranges between 0.07 and 0.12 across the years.

Selection on Observables. Less educated and less experienced workers chose to work in Israel; those with better skills chose to work locally and

[18]The table and figures report *point* estimates. The linear-quadratic experience premia *profile* in the local economy lies well above that of Israel.

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

were compensated for the baseline wage differential by the local returns given to their skills. This represents negative selection on observed skills. This sorting pattern, implied by the results of estimation, is borne out by the actual, observed locational distributions by education and age. Borjas, Kauppinen, and Poutvaara (2019) show[20] 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.

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, 19% of local workers were employed in high-skilled occupations (the top three in the table) vs. 1% 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) on returns to analytical and manual skills (see their Tables 5 and 6), using detailed task and job data on the U.S. This last point is key, as will be shown in the interpretation of the results against the background of the development accounting literature.

The phenomenon of low skill premia for movers is 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 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 was no search process for good matches and the employment relationship was not of long duration.

Selection on Unobservables. The last term in equation (21), $\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\widehat{\overline{\lambda}_{local}} - \widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\widehat{\overline{\lambda}_{Israel}}$, ranges between -0.09 and $+0.03$. It thus contributes relatively little to the explanation of the wage differential across location.

[20] See their page 150 and equation 12.

The next sub-section goes into detail about the type of selection involved here.

Accounting for the Wage Differential. The afore-going discussion paints the following picture. While there is variation across sample years, the constant in Israel is substantially higher, i.e., $\widehat{k}_{Israel} >> \widehat{k}_{local}$; the converse in true for the task component whereby $\widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} >> \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel}$. The skill premia difference, with $\widehat{\boldsymbol{\beta}}_{local} >> \widehat{\boldsymbol{\beta}}_{Israel}$, played the major role. The differences in self-selection on unobservables were relatively small. Hence the afore-cited two big components offset each other to a large extent, yielding a small wage differential (four times in favor of the Israeli location, twice in favour of the local location, and once there was no differential across sample years).

6.2 Patterns of Self-Selection on Unobservables

The discussion above has made it clear that selection on observables was negative. Self-selection on unobservables was positive as evidenced by the results in Table 3 (see the positive estimates of ρ_i). To be more specific, 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:

$$\begin{aligned}
& E\left(\ln w_i \mid \ln w_i + \ln\left[1-k_i(\mathbf{L})\right] + \ln[1-\gamma_i] > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right]\right) + \ln[1-\gamma_j] \\
= \; & \ln \pi_i + \mu_i + \frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}\lambda(c_i) \qquad (22) \\
& var\left(\ln w_i \mid \ln w_i + \ln\left[1-k_i(\mathbf{L})\right] + \ln[1-\gamma_i] > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right]\right) + \ln[1-\gamma_j] \\
= \; & \sigma_{ii}\left\{ \begin{array}{c} \rho_i^2[1-c_i\lambda(c_i)-\lambda^2(c_i)] \\ +(1-\rho_i^2) \end{array} \right\} \qquad (23)
\end{aligned}$$

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 $\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$.[21] Assuming, without loss of generality, that $\sigma_{jj} \geq \sigma_{ii}$, the

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

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

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:[22](i) positive correlation between the countries and relatively high, i.e., $\rho_{ij} \geq \frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$. Selection is positive in location j and negative in i; (ii) negative correlation between the countries i.e., $\rho_{ij} < 0$. 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., $0 \leq \rho_{ij} < \frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ and in each location there is positive 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. The results of estimation indicate that (i) the correlation $\rho_{israel,local}$ is lower than the ratio of standard deviations $\frac{\sqrt{\sigma_{israel}}}{\sqrt{\sigma_{local}}}$; in five out of the seven years it is negative; (ii) the variance in local employment is higher than that of employment in Israel ($\sigma_{local} > \sigma_{israel}$). Hence the second case (in five of the seven sample years) and the third case (in the remaining two years) above obtain, with positive self-selection in both locations.

These results are reasonable in terms of the afore-going discussion. The low positive correlation of unobserved skills across locations, or, more frequently, the negative correlation, is probably due to the fact that local and Israeli occupational tasks differed substantially, as discussed above. They are consistent with the findings of Autor and Handel (2013) on bivraiate relationships between returns on abstract, analytical and on manual skills (see their Table 7),which are also negative.

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) $\rho_{ij}\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}} > 1$. The findings here indicate that there is no stochastic dominance in unobservable skills, given the afore-cited low correlation of unobserved skills (ρ_{ij}) across locations.

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

7 Wage and Skill Distributions and The Moving Decision

Thus far the analysis was mostly in terms of means. I now look at the roles of the skills and wage *distributions* in shaping the decision to move from poor to rich economies. The moving decision can be represented graphically in 3D plots which illuminate a number of aspects.

7.1 Wages as a Function of Skills

The discussion above has yielded the following wage equations:

$$\ln w_i = \widehat{k}_i + \widehat{\boldsymbol{\beta}}_i \mathbf{X}_i + \widehat{\rho}_i \widehat{\sqrt{\sigma_i}} \widehat{\lambda}_i + v_i \tag{24}$$

Using the estimates of Table 3 for 1981 (in panel a) and for 1987 (in panel b), Figure 4 plots the following equation:

$$\ln w_i \mid \left(E\widehat{\rho}_i \widehat{\sqrt{\sigma_i}} \widehat{\lambda}_i, Ev_i \right) = C_i + \beta^{educ} educ_i + \beta_1 S_{h,i} + \beta_2 S_{h,i}^2 \tag{25}$$

where $S_{h,i}$ is experience and where

$$C_i = \widehat{k}_i + \beta_{i,0} + E\left(\widehat{\rho}_i \widehat{\sqrt{\sigma_i}} \widehat{\lambda}_i \right)$$

and v_i is set to zero.

This is a graph of (conditional) log wages as a function of education and experience, holding constant the selection term $E\left(\widehat{\rho}_i \widehat{\sqrt{\sigma_i}} \widehat{\lambda}_i \right)$, taking into account $\widehat{k}_i + \beta_{i,0}$, without the unobserved skills. So it is a 3D plot with log wages, education and experience on the axes. It gives expression to the distributions of wages and skills, not just to the mean points.

Figure 3

The graphs show that at low levels of education and experience, the blue plane, representing log wages of workers in Israel lies above the green plane, representing log wages in local employment. Workers choose employment in Israel, conditional on C_i and $v_i = 0$. The positions are reversed at high levels of education and experience, where workers choose local employment. The demarcation line, where the switch occurs, is denoted by the dashed line. This, then, is a depiction of the negative selection on education and experience, whereby the low (high) skilled workers choose work in Israel (locally) where for them wages are higher.

To get a sense of the magnitudes embodied in Figure 4, Table 7 shows log wages for each location as predicted by equation (25) as well as the part predicted by average skills in the locality. It does so for 1981 and for 1987.

Table 7

The values of the first row in the table are marked (by I and L, for Israel and local) in Figure 4. The table shows that local workers are more skilled relative to workers in Israel (see the two bottom rows). These workers would get much higher wages due to the skills differences and the differences in returns on them (second row). In 1981 this skill-induced difference in log wage terms was 0.96 log points and in 1987 it was 0.61 log points. But overall workers in Israel got higher wages (first row) because of the other terms in the equation, contained in C_i and discussed above. These differences in favor of workers in Israel amount to 0.06 log points in 1981 and 0.15 log points in 1987.

7.2 Tasks and Unobserved Skills

A different angle is provided by looking at the relation between tasks in the two locations. The analysis above yields:[23]

$$\begin{aligned}\ln t_{Israel} &= \mu_{Israel} + \frac{\sigma_{local,Israel}}{\sigma_{local}}(\ln t_{local} - \mu_{local}) + \varepsilon_{Israel} \\ &= \left(\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}\right) + \frac{\sigma_{local,Israel}}{\sigma_{local}} \ln t_{local} + \varepsilon_{Israel}\end{aligned} \quad (26)$$

where:

$$\begin{aligned}\varepsilon_{Israel} &= u_{Israel} - u_{local}\frac{\sigma_{local,Israel}}{\sigma_{local}} \\ E\varepsilon_{Israel} &= 0 \\ var\ \varepsilon_{Israel} &= \sigma_{Israel}[1 - \frac{\sigma^2_{local,Israel}}{\sigma_{local}\sigma_{Israel}}]\end{aligned}$$

Figure 4 depicts this relation in the 3D space of log tasks ($\ln t_{local}, \ln t_{Israel}$) and ε_{Israel} (expressing adjusted differences between unobserved skills in Israel and in the local economy), using the point estimates and second moments for 1981 and for 1987 from Table 3 (in two panels).

Figure 4

To understand the figure note the following elements. For any given worker, his log task value in each location is indicated on two axes and his adjusted unobserved skills differences (ε_{Israel}) value is given on the third

[23] Derived from multiplying both sides of the equation $\ln t_{local} = \mu_{local} + u_{local}$ by $\frac{\sigma_{local,Israel}}{\sigma_{local}}$ and subtracting from $\ln t_{Israel}$.

axis. 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 $\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}$,[24] 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. The data points are distributed – conditional on the $\ln t_{local}$ value – with $var\ \varepsilon_{Israel}$. The other (black) line in the figure is the 45 degree line serving as the line of equal income ($\ln w_{local} = \ln w_{Israel}$).[25]This 45 degree line is the demarcation line in this figure 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.

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.[26]

Negative selection on observables. Moving along the (red) regression line, the workers with relatively low $\ln t_{local}$ (low observable skills) 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 (those with high observable skills) choose to work locally. This is also what was seen in the depiction of the wage-skills relations in Figure 4.

Positive selection on unobservables. The figure illustrates the positive selection on unobservables in each location.[27]In 1981, the term $\frac{\sigma_{local,Israel}}{\sigma_{local}}$ is positive and less than 1. The regression line is less steep than the black 45 degrees line and starts above it. In 1987, as in most of the sample years, the regression slope is negative. Thus, in both cases, 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 in the normal distribution 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.

One question of interest is to consider how moving behavior would

[24]I use the point estimates of the coefficients (from Table 3) in 1981 and 1987, and the sample means of the X variables, to generate μ_{local} and μ_{Israel}. I adopt the normalization of $\beta_0 = 0$.

[25]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 $\ln t_j = \ln \pi_i - \ln \pi_j + \ln t_i$.

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

[27]In terms of equation (**??**) this means that in each sector

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

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 5 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 Implications for Development Accounting

Following the discussion of the literature in sub-section 2.1, the derivation discussed in sub-section 3.2 and in Appendix A, and using equation (17), productivity differences across locations are given by:

$$\ln z_i - \ln z_j = \ln \pi_i - \ln \pi_j \tag{27}$$

The estimates of Table 5 relate to $\widehat{k}_i = \ln \widehat{\pi}_i + \widehat{\beta_{i,0}}$. The presence of the task function intercept makes the estimated $\widehat{k}_i - \widehat{k}_j$ a lower bound on task prices π or productivity z differentials. The estimates of $\widehat{k}_i - \widehat{k}_j$ vary between 0.48 and 1.09 log points (across the seven years of repeated cross sections) in favor of the Israeli economy. This implies a lower bound on the $\frac{z_{Israel}}{z_{local}}$ ratio ranging between 1.6 and 3. Hendricks and Schoellman (2018) report z ratios of very similar magnitude.

Note that the total wage differential, ranging between -0.08 and $+0.17$ log points, masks this substantial gain in productivity.

This finding is important in the context of tying the development accounting analysis to the current analysis. In this context, the results here are basically as follows:

(i) TFP and capital stock differences are large; there are substantial productivity differences (the z_i) in favor of the rich economy (Israel), 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 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), selection on unobservables is not very important quantitatively.

The skill premia differences discussed in point (ii) offset (partially or completely) the wage gains of movers, and contradict the claims of huge mover wage gains. The issue is that the human capital embodied in the movers operates to lower wages as they come with relatively low skills and are demanded in particular job tasks, which require such low skills.

There are important differences between the results of the afore-cited HS (2018, 2019) studies, which are key in the DA literature, and the results of this paper.

First, the empirical objects studied are not the same. HS use wage data from the New Immigrant Survey, a representative sample of adult immigrants granted lawful permanent residence in the United States ("green card" recipients) between May and November 2003, drawn from government administrative records. They have data on up to two pre-migration jobs and up to three post-migration jobs and do PPP adjustments. Hence they look at the same workers pre- and post-migration. The current paper does not compare wages of the same workers across countries as HS do (home and the U.S) but rather looks at wages of movers and stayers. Data are taken from repeated cross-sections in the period 1981-1987 and pertain to current year jobs. HS (2018, 2019) use GDP per worker across countries and deduce estimates of z differences from comparing GDP per worker to pre- and post-migration wage differences. Their z differences relate to GDP. The current paper does not use any output data, and derives estimates of z differences from wage regressions across locations, with wages relating to output in the relevant jobs. Thus z differences here relate to locational/sectorial output.

Second, the HS (2018,2019) computations, using GDP per capita and pre- and post- migration wage data, assume, in the baseline scenario, that human capital is fully transferable and are 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 aforecited wages, across the U.S. and source countries (see their equation 4 in HS (2018)). HS (2018) find that wage differentials are much lower than GDP per capita differentials (their Table II). Thus, they reach the conclusion that human capital differences play a big role, between 48% and 66%, in cross country income differences. This conclusion remains broadly true when taking into account a whole host of factors, in particular, imperfect substitution of skills in production and endogenous choice of production (see Tables IV,V and VIII in HS (2018) and Tables 7, 10, 11, 13, and 14 in HS (2019)). The current paper has tasks, rather than human capital stocks per se, in producing a good in a particular location, and the latter production is not GDP. Tasks are defined by location and are bundles of skills, with returns to these skills included. 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 (14)) of a location reflect the country. The task bundle reflects the worker (his skills, $\mathbf{X}$ and his job task returns ($\boldsymbol{\beta}$)).

From the above it is clear that the wage differential examined here, across locations, is not the same as the HS wage differential. The wage differential here reflects both the z cross-country differential, as in HS,[28] *as well as* the task differential across locations, which reflects worker skills and job task returns, *unlike* the approach of HS.

The HS results may still hold true in the current case. Human capital is higher in Israel and it is highly likely that human capital differences play a big role in the GDP per capita differential, which is a factor of about 5 and more here. These points, however, are not examined in the current paper. Likewise, the findings here, whereby the foreign task bundle has low value in terms of wages for the movers, is not an issue examined by HS. The HS papers do not study the task composition of pre- and post-migration jobs. The low task value found here is consistent with both the HS view on lower human capital in poor countries, and the findings, related to human capital in poor countries of Lagakos et al (2018a,b). Thus, large differences in human capital explain the offset effect here, through task values, which lowers the wages of movers.

It should be noted that HS (2018) also examine the effects of selection. They find positive selection on observables and on unobservables (see their

[28]The discussion above shows that the results here are of very similar magnitude (relative to HS (2018)) on this dimension.

Figure II) while the current paper finds negative selection on observables, which plays an important role quantitatively, and positive selection on unobservables. They find evidence in favor of gaps in the marginal value product of labor across sectors. These gaps imply that each country's aggregate z and average wage gains at migration are affected by the sectoral composition of employment. This last point is inherent in the analysis undertaken here.

A key emerging insight is the following. The task-based model of Roy (1951), further developed by Heckman and Sedlacek (1985), which has received so much application in migration studies posits that workers choose locations which are related to the performance of particular tasks. Thus, movers are not performing the same tasks in the home and host countries. This has important repercussions in terms of the rewards to skills which they get and, as the analysis here demonstrates, in terms of their wage gains. The afore-cited empirical analysis of Autor and Handel (2013) which explicitly examines wages, jobs, and tasks within the framework of this model, is of particular importance. It indicates that this model is an empirically-relevant one.

9 Conclusions

The move from poor to rich countries is a prevalent and important phenomenon; recent literature has emphasized the large potential gains inherent in it. It ties in with the important current discussion in the development accounting literature. This paper exploits a case which facilitates the study of this move without confounding factors. It turns out that the substantial gross productivity and human capital differences across rich and poor economies play opposing roles, yielding much lower net gains. The 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, and potentially contradictory, effects.

While 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 analysis here, in disentangling the different effects at play, contributes the following key lessons.

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

(ii) These 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 not gain from the human capital differentials across countries, as they stay with their poor country skills, and are constrained by the task requirements they face.

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

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 human capital differences, within a task-based approach.

Recent literature (see, for example, Acemoglu and Autor (2011, pp. 1070-1096), Autor and Salomons (2018), and Acemoglu and Restrepo (2019)) has shown that there are changes in productivity, wage, and occupational distributions related to changing tasks distributions. Technological processes, like increased automation and the related decline in routine jobs, change task requirements in significant ways. These processes are pertinent in the current context. It is highly plausible to imagine that foreign and home tasks requirements undergo changes, and so task requirements of movers and stayers will change. Hence a task-based approach is crucial in terms of an empirically-relevant model of the move from a poor to a rich economy.

References

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

[2] Acemoglu, Daron and Pascual Restrepo, 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor," **Journal of Economic Perspectives** 33, 2, 3–30.

[3] Acemoglu, Daron and David H. Autor, 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Orley Ashenfelter and David E. Card (eds.), **Handbook of Labor Economics** Volume 4, 1043–1171, Amsterdam: Elsevier.

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

[5] Arnon, Arie, Israel Luski, Avia Spivak, and Jimmy Weinblatt, 1997. T**he Palestinian Economy.** Leiden, New York and Koln, Brill.

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

[7] Autor, David H. and Anna Salomons, 2018. "Is Automation Labor Share–Displacing? Productivity Growth, Employment, and the Labor Share," **Brookings Papers on Economic Activity**, Spring, 1-63.

[8] Banerjee, Abhijit V. and Benjamin Moll, 2010."Why Does Misallocation Persist?," **American Economic Journal: Macroeconomics** 2:1, 189–206.

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

[10] 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.

[11] Bryan, Gharad and Melanie Morten, 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," **Journal of Political Economy** 127,5, 2229 - 2268.

[12] Caselli, Francesco, 2005. "Accounting for Cross-Country Income Differences. In: **Handbook of Economic Growth**, vol. 1, pp. 679–741, North Holland.

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

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

[15] 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.

[16] D'Haultfoeuille, X., A. Maurel, and Y. Zhang, 2018. "Extremal Quantile Regressions for Selection Models and the Black -White Wage Gap." **Journal of Econometrics** 203(1): 129-142.

[17] D'Haultfoeuille, X., A. Maurel, X. Qiu, and Y. Zhang, 2019. "Estimating Selection Models without Instrument with Stata," NBER Working Paper No. 25823.

[18] Dustmann,Christian and Tommaso Frattini, 2013. "Immigration: The European Experience" in David Card and Steve Raphael (eds.) **Immigration, Poverty, and Socioeconomic Inequality**, Chapter 13, 423-456, New York, NY: Russel Sage Foundation.

[19] Dustmann,Christian, Tommaso Frattini, and Ian P. Preston, 2013. "The Effect of Immigration along the Distribution of Wages," **Review of Economic Studies** 80, 145–173

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

[21] Dustmann, Christian and Ian P. Preston, 2019. "Free Movement, Open Borders, and the Global Gains from Labor Mobility," **Annual Review of Economics** 11, 783-808.

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

[23] 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.

[24] 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.

[25] Hendricks, Lutz and Todd Schoellman, 2019. "Skilled Labor Productivity and Cross-country Income Differences," working paper.

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

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

[28] 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.

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

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

[31] 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.

[32] 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.

[33] Peri, Giovanni , 2016. "Immigrants, Productivity, and Labor Markets," **Journal of Economic Perspectives** 30,4, 3–30.

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

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

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

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

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

10 Tables

Table 1
Sample Statistics
Palestinian Male Workers, 1981-1987

	1981		1982		1983		1984	
	Local	**Israel**	**Local**	**Israel**	**Local**	**Israel**	**Local**	**Israel**
N	$5,370$	$7,345$	$5,402$	$7,715$	$5,328$	$8,165$	$5,666$	$8,772$
log wage	-4.54	-4.51	-3.73	-3.69	-2.84	-2.80	-1.33	-1.50
(hourly)	(0.61)	(0.46)	(0.57)	(0.46)	(0.63)	(0.47)	(0.77)	(0.64)
education	7.69	6.34	7.93	6.63	8.24	6.87	8.45	7.05
(years)	(4.82)	(3.95)	(4.90)	(3.91)	(4.82)	(3.87)	(4.81)	(3.93)
experience	21.78	20.61	21.42	19.90	20.73	19.34	20.14	19.08
(years)	(14.80)	(14.74)	(14.36)	(14.33)	(14.23)	(14.19)	(14.13)	(14.21)
residence								
Jenin	0.09	0.08	0.07	0.08	0.08	0.09	0.08	0.09
Nablus	0.19	0.05	0.18	0.05	0.19	0.06	0.18	0.05
Tulkarm	0.09	0.12	0.08	0.11	0.08	0.12	0.09	0.13
Ramallah	0.15	0.13	0.14	0.14	0.16	0.14	0.17	0.13
Jordan valley	0.03	0.00	0.03	0.01	0.03	0.00	0.02	0.01
Bethlehem	0.10	0.08	0.10	0.08	0.10	0.07	0.10	0.09
Hebron	0.16	0.18	0.17	0.17	0.17	0.17	0.16	0.16
Rafiah	0.02	0.05	0.01	0.04	0.01	0.03	0.02	0.04
Gaza	0.14	0.21	0.16	0.21	0.14	0.20	0.14	0.19
Khan Yunis	0.05	0.11	0.06	0.11	0.05	0.11	0.04	0.11
rural	0.36	0.52	0.34	0.52	0.36	0.53	0.37	0.54
urban	0.48	0.24	0.50	0.25	0.49	0.24	0.48	0.24
refugee camp	0.16	0.24	0.16	0.23	0.14	0.23	0.15	0.21
married	0.74	0.76	0.74	0.73	0.72	0.72	0.71	0.71

	1985		1986		1987	
	Local	**Israel**	**Local**	**Israel**	**Local**	**Israel**
N	$6,111$	$8,812$	$6,835$	$9,607$	$7,250$	$11,582$
log wage	0.08	-0.06	0.64	0.64	0.90	0.97
(hourly)	(0.62)	(0.55)	(0.49)	(0.41)	(0.44)	(0.36)
education	8.43	7.22	8.70	7.49	8.93	7.73
(years)	(4.72)	(3.92)	(4.65)	(3.93)	(4.54)	(3.88)
experience	19.63	18.61	18.98	17.99	18.49	17.55
(years)	(14.06)	(13.81)	(13.59)	(13.47)	(13.11)	(13.23)
residence						
Jenin	0.09	0.08	0.08	0.08	0.08	0.10
Nablus	0.17	0.05	0.17	0.06	0.17	0.06
Tulkarm	0.08	0.13	0.07	0.14	0.07	0.14
Ramallah	0.16	0.13	0.16	0.13	0.17	0.13
Jordan valley	0.02	0.01	0.02	0.01	0.02	0.01
Bethlehem	0.09	0.10	0.10	0.09	0.11	0.12
Hebron	0.18	0.16	0.19	0.15	0.20	0.17
Rafiah	0.02	0.05	0.01	0.05	0.02	0.03
Gaza	0.15	0.19	0.15	0.19	0.13	0.15
Khan Yunis	0.04	0.11	0.04	0.10	0.04	0.09
rural	0.38	0.54	0.37	0.55	0.41	0.62
urban	0.48	0.25	0.50	0.25	0.47	0.22
refugee camp	0.14	0.21	0.14	0.21	0.12	0.17
married	0.69	0.70	0.66	0.69	0.68	0.67

Notes:

1. The wage distribution was truncated at 1% at the bottom and at 0.2% at the top of the distribution.

2. For log wages, years of education and years of experience, the table reports the mean of variables with standard deviations in parentheses.

3. The region of residence, type of residence and married numbers are percentage of workers out of total sample in the column.

Table 2: Heckman Two Step Estimation 1987

**a. The Selection Equation:
Probability of selection of employment in Israel**

	1	2
constant	0.54***	1.37***
	(0.096)	(0.102)
education	−0.09***	−0.09***
	(0.003)	(0.003)
experience	−0.03***	−0.04***
	(0.003)	(0.004)
experience^2/100	0.03***	0.04***
	(0.005)	(0.006)
married		0.17***
		(0.030)
urban residence		−0.99***
		(0.026)
refugee camp residence		−0.36***
		(0.032)
Jenin	1.00***	0.35***
Nablus	0.24***	−0.17*
Tulkarm	1.30***	0.83***
Ramallah	0.70***	0.08
Bethlehem	0.93***	0.42***
Hebron	0.71***	0.24***
Rafiah	1.32***	1.13***
Gaza	0.97***	0.96***
Khan Yunis	1.46***	1.22***

b. The Wage Regression

	(1)		(2)		(3)	
exclusion	**one, Set 1**		**three, Set 2**		**OLS**	
restrictions	Local	Israel	Local	Israel	Local	Israel
constant	-0.125^{**}	0.582^{***}	0.021	0.583^{***}	0.110^{***}	0.583^{***}
	(0.040)	(0.017)	(0.027)	(0.017)	(0.020)	(0.017)
Q2	0.073^{***}	0.113^{***}	0.079^{***}	0.112^{***}	0.080^{***}	0.112^{***}
	(0.013)	(0.009)	(0.013)	(0.009)	(0.013)	(0.009)
Q3	0.055^{***}	0.178^{***}	0.068^{***}	0.177^{***}	0.068^{***}	0.177^{***}
	(0.014)	(0.009)	(0.013)	(0.009)	(0.013)	(0.009)
Q4	0.139^{***}	0.246^{***}	0.145^{***}	0.246^{***}	0.144^{***}	0.246^{***}
	(0.013)	(0.009)	(0.013)	(0.009)	(0.013)	(0.009)
education	0.044^{***}	0.010^{***}	0.039^{***}	0.012^{***}	0.037^{***}	0.012^{***}
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
experience	0.036^{***}	0.017^{***}	0.034^{***}	0.017^{***}	0.033^{***}	0.017^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
experience^2	-0.047^{***}	-0.027^{***}	-0.045^{***}	-0.028^{***}	-0.044^{***}	-0.028^{***}
(/100)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
λ	0.150	0.029	0.063	0.002		
ρ_i	0.362	0.084	0.157	0.004		
$\sqrt{\sigma_{ii}}$	0.415	0.346	0.401	0.345		
R^2					0.187	0.094
Wald/F test	$1,335$	$1,131$	$1,576$	$1,144$	278	200
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
n	$7,248$	$11,580$	$7,248$	$11,580$	$7,248$	$11,580$

Notes:

1.The equation in panel a relates to the probability of selection of employment in Israel. The specifications are elaborated in Section 3; see, in particular, equation (12).

2. The wage equation in panel b is estimated with two sets of exclusion restrictions in columns 1 and 2, respectively, and uses OLS in column 3. For the exclusion restrictions, Set 1 is given by

$$\begin{aligned}\mathbf{L} &\in [\text{region of residence}] \\ \mathbf{X} &\in [\text{education, experience}]\end{aligned}$$

Set 2 is given by

$$\begin{aligned}\mathbf{L} &\in [\text{region of residence, marital status, urban status }] \\ \mathbf{X} &\in [\text{education, experience}]\end{aligned}$$

3. The sample includes all wage earners except those with hourly wages below the lowest 1% or above the highest 0.2%.

4. Standard errors of the coefficients are in parentheses, except for the region of residence variables in panel a.

5. Three stars denote significance at 1%, two at 5%, and one at 10%.

6. The baseline region of residence is the Jordan valley and the baseline type of residence is rural.

7. The second moments satisfy the following relation:

$$\rho_i = \frac{\sigma_{ii} - \sigma_{ij}}{\sqrt{\sigma_{ii}}\sigma^*}$$

Table 3: Heckman Two Step Estimation
1981-1987

a. The selection equation

	1981	1982	1983	1984	1985	1986	1987
constant	0.084	0.509***	0.217	0.837***	0.493***	0.515***	0.543***
	(0.128)	(0.120)	(0.133)	(0.110)	(0.115)	(0.111)	(0.096)
education	−0.095***	−0.093***	−0.095***	−0.093***	−0.089***	−0.086***	−0.087*
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
experience	−0.028***	−0.033***	−0.036***	−0.035***	−0.027***	−0.030***	−0.033*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
exp^2	0.013*	0.022***	0.026***	0.027***	0.013*	0.018***	0.025***
/100	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Jenin	1.329***	1.028***	1.442***	0.781***	0.823***	0.796***	0.997***
Nablus	0.530***	0.121	0.524***	−0.144	0.059	0.106	0.239**
Tulkarm	1.574***	1.211***	1.606***	0.928***	1.189***	1.223***	1.304***
Ramal.	1.245***	0.973***	1.231***	0.546***	0.686***	0.603***	0.700***
Beth.	1.255***	0.901***	1.126***	0.655***	0.976***	0.786***	0.933***
Hebron	1.326***	0.903***	1.221***	0.610***	0.723***	0.597***	0.713***
Rafiah	2.041***	1.656***	1.955***	1.219***	1.621***	1.611***	1.319***
Gaza	1.593***	1.158***	1.573***	0.866***	1.016***	0.973***	0.973***
K. Yunis	1.762***	1.398***	1.828***	1.356***	1.487***	1.377***	1.460***

b. The Wage Equation

	(1) Local1981	(2) Israel1981	(3) Local1982	(4) Israel1982	(5) Local1983	(6) Israel1983
Constant	-6.236^{***}	-5.171^{***}	-5.371^{***}	-4.309^{***}	-4.274^{***}	-3.424^{***}
	(0.055)	(0.024)	(0.052)	(0.022)	(0.053)	(0.022)
Q2	0.279^{***}	0.264^{***}	0.237^{***}	0.236^{***}	0.186^{***}	0.207^{***}
	(0.020)	(0.013)	(0.018)	(0.012)	(0.020)	(0.013)
Q3	0.480^{***}	0.471^{***}	0.416^{***}	0.455^{***}	0.445^{***}	0.462^{***}
	(0.020)	(0.013)	(0.019)	(0.012)	(0.020)	(0.012)
Q4	0.660^{***}	0.632^{***}	0.624^{***}	0.698^{***}	0.748^{***}	0.721^{***}
	(0.020)	(0.013)	(0.018)	(0.012)	(0.020)	(0.012)
education	0.070^{***}	0.011^{***}	0.064^{***}	$0.006 **$	0.055^{***}	0.006^{**}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
experience	0.044^{***}	0.017^{***}	0.043^{***}	0.013^{***}	0.044^{***}	0.014^{***}
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
experience^2/100	-0.050^{***}	-0.029^{***}	-0.050^{***}	-0.024^{***}	-0.058^{***}	-0.025^{***}
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
λ	0.229	0.148	0.271	0.201	0.095	0.192
ρ_i	0.431	0.369	0.536	0.508	0.183	0.479
$\sqrt{\sigma_{ii}}$	0.530	0.403	0.505	0.396	0.517	0.402
$\frac{\sqrt{\sigma_{Israel}}}{\sqrt{\sigma_{local}}}$	0.759		0.784		0.776	
$\rho_{Israel,local}$	0.273		0.640		−1.786	
Wald Test	2,305	2,921	2,360	3,684	2,537	4,021
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	5,368	7,337	5,401	7,711	5,328	8,165

	(7)	(8)
	Local1984	Israel1984
Constant	-3.203^{***}	-2.508^{***}
	(0.054)	(0.026)
Q2	0.470^{***}	0.418^{***}
	(0.020)	(0.014)
Q3	0.946^{***}	0.882^{***}
	(0.020)	(0.014)
Q4	1.307^{***}	1.203^{***}
	(0.019)	(0.014)
education	0.063^{***}	0.007^{**}
	(0.002)	(0.002)
experience	0.044^{***}	0.017^{***}
	(0.002)	(0.002)
experience^2/100	-0.054^{***}	-0.030^{***}
	(0.003)	(0.003)
λ	0.079	0.315
ρ_i	0.151	0.639
$\sqrt{\sigma_{ii}}$	0.523	0.493
$\frac{\sqrt{\sigma_{Israel}}}{\sqrt{\sigma_{local}}}$	0.941	
$\rho_{Israel,local}$	-1.104	
Wald Test	6,043	8,130
p-value	(0.000)	(0.000)
N	5,666	8,771

	(9) Local1985	(10) Israel1985	(11) Local1986	(12) Israel1986	(13) Local1987	(14) Israel1987
Constant	-1.317^{***}	-0.878^{***}	-0.384^{***}	0.210^{***}	-0.125^{**}	0.582^{***}
	(0.047)	(0.025)	(0.040)	(0.022)	(0.040)	(0.017)
Q2	0.345^{***}	0.348^{***}	0.081^{***}	0.133^{***}	0.073^{***}	0.113^{***}
	(0.018)	(0.013)	(0.015)	(0.012)	(0.013)	(0.009)
Q3	0.634^{***}	0.725^{***}	0.170^{***}	0.246^{***}	0.055^{***}	0.178^{***}
	(0.018)	(0.014)	(0.015)	(0.012)	(0.014)	(0.009)
Q4	0.757^{***}	0.827^{***}	0.219^{***}	0.253^{***}	0.139^{***}	0.246^{***}
	(0.017)	(0.014)	(0.015)	(0.012)	(0.013)	(0.009)
education	0.053^{***}	$0.006 **$	0.047^{***}	$0.004*$	0.044^{***}	0.010^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
experience	0.041^{***}	0.017^{***}	0.038^{***}	0.014^{***}	0.036^{***}	0.017^{***}
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\text{experience}^2/100$	-0.050^{***}	-0.029^{***}	-0.048^{***}	-0.026^{***}	-0.047^{***}	-0.027^{***}
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
λ	0.018	0.254	0.046	0.183	0.150	0.029
ρ_i	0.037	0.544	0.106	0.444	0.362	0.084
$\sqrt{\sigma_{ii}}$	0.492	0.467	0.433	0.411	0.415	0.346
$\frac{\sqrt{\sigma_{Israel}}}{\sqrt{\sigma_{local}}}$	0.949		0.949		0.833	
$\rho_{Israel,local}$	-1.061		-1.089		-0.744	
Wald Test	3,262	4,865	1,507	796	1,335	1,131
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	6,110	8,812	6,833	9,603	7,248	11,580

Notes:
1. See notes 1, 3-6 in Table 1.
2. Panel b uses set 1 for the exclusion restrictions given by

$$\begin{aligned}
\mathbf{L} &\in [\text{region of residence}] \\
\mathbf{X} &\in [\text{education, experience}]
\end{aligned}$$

The first stage reported in panel a.
3. The second moments satisfy the following relations:

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

Hence:

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

Solving the last equation for σ_{ij} (using $\rho_i, \rho_j, \sqrt{\sigma_{jj}}, \sqrt{\sigma_{jj}}$) the cross location correlation $\rho_{ij} \equiv \rho_{Israel,local}$ is computed.

Table 4 : Heckman and Semi Parametric Estimation
1981-1987

1981	**Semi Parametric**			**Heckman**		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.052***	0.021***	−0.031	0.070***	0.011***	−0.059
	(0.003)	(0.004)		(0.002)	(0.002)	
exp	0.022***	0.022***	0.000	0.044***	0.017***	−0.027
	(0.004)	(0.003)		(0.002)	(0.001)	
exp^2	−0.021***	−0.032***	−0.011	−0.050***	−0.029***	0.020
/100	(0.006)	(0.006)		(0.003)	(0.002)	

1982	**Semi Parametric**			**Heckman**		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.041***	0.018***	−0.023	0.064***	0.006***	−0.058
	(0.002)	(0.003)		(0.002)	(0.002)	
exp	0.023***	0.019***	−0.004	0.043***	0.013***	−0.030
	(0.002)	(0.003)		(0.002)	(0.001)	
exp^2	−0.024***	−0.026***	−0.001	−0.050***	−0.024***	0.026
/100	(0.005)	(0.005)		(0.003)	(0.002)	

1983	**Semi Parametric**			**Heckman**		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.044***	0.013***	−0.031	0.055***	0.006***	−0.049
	(0.002)	(0.003)		(0.002)	(0.002)	
exp	0.030***	0.015***	−0.016	0.044***	0.014***	−0.031
	(0.002)	(0.002)		(0.002)	(0.001)	
exp^2	−0.036***	−0.023***	0.013	−0.058***	−0.025***	0.033
/100	(0.004)	(0.003)		(0.004)	(0.002)	

1984	**Semi Parametric**			**Heckman**		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.051***	0.030***	−0.021	0.063***	0.007***	−0.056
	(0.002)	(0.003)		(0.002)	(0.002)	
exp	0.031***	0.027***	−0.004	0.044***	0.017***	−0.027
	(0.002)	(0.003)		(0.002)	(0.002)	
exp^2	−0.034***	−0.037***	−0.002	−0.054***	−0.030***	0.024
/100	(0.004)	(0.006)		(0.003)	(0.003)	

1985	Semi Parametric			Heckman		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.051^{***}	0.020^{***}	-0.031	0.053^{***}	0.006^{***}	-0.047
	(0.003)	(0.003)		(0.002)	(0.002)	
exp	0.029^{***}	0.024^{***}	-0.006	0.041^{***}	0.017^{***}	-0.024
	(0.003)	(0.002)		(0.002)	(0.001)	
$\exp^2$	-0.029^{***}	-0.035^{***}	-0.006	-0.050^{***}	-0.029^{***}	0.021
/100	(0.006)	(0.003)		(0.003)	(0.002)	

1986	Semi Parametric			Heckman		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.040^{***}	0.014^{***}	-0.026	0.047^{***}	$0.004**$	-0.043
	(0.002)	(0.003)		(0.002)	(0.002)	
exp	0.027^{***}	0.015^{***}	-0.012	0.038^{***}	0.014^{***}	-0.024
	(0.002)	(0.002)		(0.001)	(0.001)	
$\exp^2$	-0.031^{***}	-0.022^{***}	0.009	-0.048^{***}	-0.026^{***}	0.022
/100	(0.003)	(0.003)		(0.003)	(0.002)	

1987	Semi Parametric			Heckman		
	local	**Israel**	**diff**	**local**	**Israel**	**diff**
educ	0.032^{***}	0.011^{***}	-0.021	0.044^{***}	0.010^{***}	-0.033
	(0.002)	(0.001)		(0.002)	(0.001)	
exp	0.026^{***}	0.014^{***}	-0.012	0.036^{***}	0.017^{***}	-0.020
	(0.002)	(0.001)		(0.001)	(0.001)	
$\exp^2$	-0.034^{***}	-0.022^{***}	0.012	-0.047^{***}	-0.027^{***}	0.020
/100	(0.004)	(0.002)		(0.003)	(0.002)	

Notes:

1. Heckman estimates are taken from Table 2.

2.The semi-parametric estimation methodology is described in sub-section 5.2.2 and in Appendix B.

Table 5
Decomposition of Mean Wages and of the Mean Wage Differential

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

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

1981	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	−4.54	−4.51	−0.03
$\widehat{k}$	−5.88	−4.83	−1.05
$\widehat{\boldsymbol{\beta}}\overline{\mathbf{X}}$	1.26	0.30	0.96
$\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$			0.84
$\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.12
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\overline{\lambda}}$	0.05	0.02	0.03

1982	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	−3.73	−3.69	−0.04
$\widehat{k}$	−5.05	−3.96	−1.09
$\widehat{\boldsymbol{\beta}}\overline{\mathbf{X}}$	1.20	0.20	1.00
$\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$			0.88
$\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.12
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\overline{\lambda}}$	0.07	0.04	0.03

1983	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	-2.84	-2.80	-0.04
$\widehat{k}$	-3.93	-3.08	-0.85
$\widehat{\beta}\overline{\mathbf{X}}$	1.12	0.22	0.90
$\overline{\mathbf{X}}_{Israel}(\widehat{\beta}_{local} - \widehat{\beta}_{Israel})$			0.79
$\widehat{\beta}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.10
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\overline{\lambda}}$	0.01	0.04	-0.03

1984	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	-1.33	-1.50	0.17
$\widehat{k}$	-2.52	-1.88	-0.64
$\widehat{\beta}\overline{\mathbf{X}}$	1.20	0.26	0.93
$\overline{\mathbf{X}}_{Israel}(\widehat{\beta}_{local} - \widehat{\beta}_{Israel})$			0.82
$\widehat{\beta}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.11
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\overline{\lambda}}$	0.01	0.10	-0.09

1985	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	0.08	-0.06	0.14
$\widehat{k}$	-0.88	-0.40	-0.48
$\widehat{\beta}\overline{\mathbf{X}}$	1.06	0.26	0.80
$\overline{\mathbf{X}}_{Israel}(\widehat{\beta}_{local} - \widehat{\beta}_{Israel})$			0.71
$\widehat{\beta}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.09
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\overline{\lambda}}$	0.00	0.07	-0.06

1986	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	0.64	0.64	0.00
$\widehat{k}$	-0.27	0.37	-0.64
$\widehat{\boldsymbol{\beta}}\overline{\mathbf{X}}$	0.96	0.20	0.76
$\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$			0.68
$\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.08
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\widehat{\lambda}}$	0.00	0.03	-0.03

1987	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	0.90	0.97	-0.08
$\widehat{k}$	-0.06	0.72	-0.78
$\widehat{\boldsymbol{\beta}}\overline{\mathbf{X}}$	0.90	0.29	0.61
$\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$			0.54
$\widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.07
$\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\widehat{\lambda}}$	0.02	0.00	0.02

Notes:
The table is based on the point estimates reported in Table 2.

Table 6
Industry and Occupation Distributions by Work Locations, 1987

a. Industry Distributions

industry	Local	Israel
agriculture	4%	12%
manufacturing	25%	20%
construction	22%	49%
commerce	6%	9%
government	32%	6%
transportation	6%	2%
personal services	5%	3%
finance	1%	0%

b. Occupation Distributions

occupation	Local	Israel
academic	6%	0%
professionals	12%	1%
managers	1%	0%
clerical workers	9%	1%
agents, sales and service	12%	14%
skilled job in agriculture	4%	13%
manufacturing and construction skilled jobs	35%	29%
unskilled	22%	42%

Note:
The sample dates from 1987 and is the same as the one used in Table 1.

Table 7
Skills and Log Wages Across Locations

	1981		1987	
	local	**Israel**	**local**	**Israel**
predicted $\ln w_i$	-4.57	-4.51	0.86	1.01
$\beta^{educ}\overline{educ_i} + \beta_1\overline{S_{h,i}} + \beta_2\overline{S}^2_{h,i}$	1.26	0.30	0.90	0.29
mean education ($\overline{educ_i}$)	7.69	6.34	8.93	7.73
mean experience ($\overline{S_{h,i}}$)	21.78	20.61	18.49	17.55

Notes:

1. Education and experience means are taken from the data.

b. Total log wages, first row, are predicted from the equation, evalauted at mean skills.

$$\ln w_i \mid \left(E\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\widehat{\lambda}_i, Ev_i \right) = C_i + \beta^{educ}educ_i + \beta_1 S_{h,i} + \beta_2 S^2_{h,i} \tag{28}$$

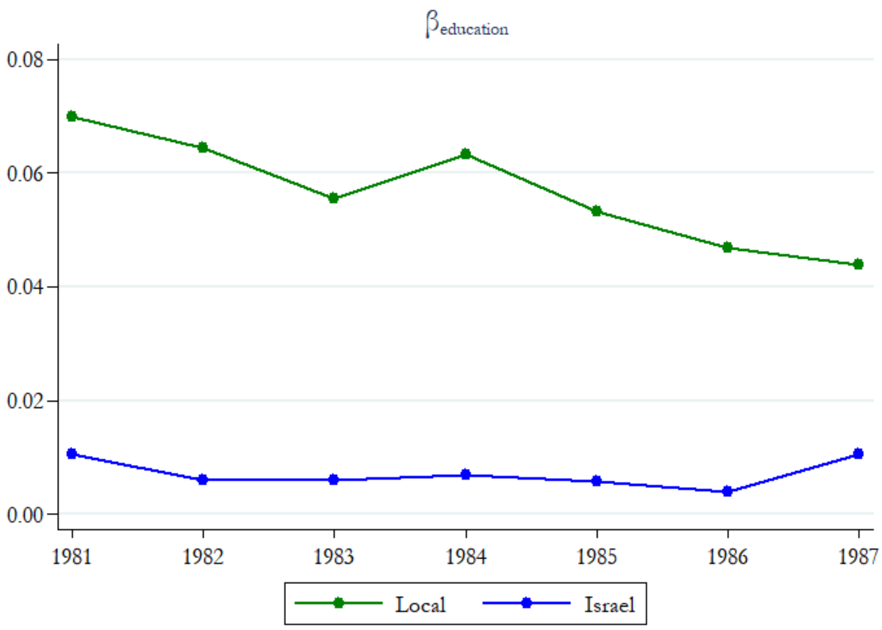
where $S_{h,i}$ is experience and where

$$C_i = \widehat{k}_i + \beta_{i,0} + E\left(\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\widehat{\lambda}_i\right)$$

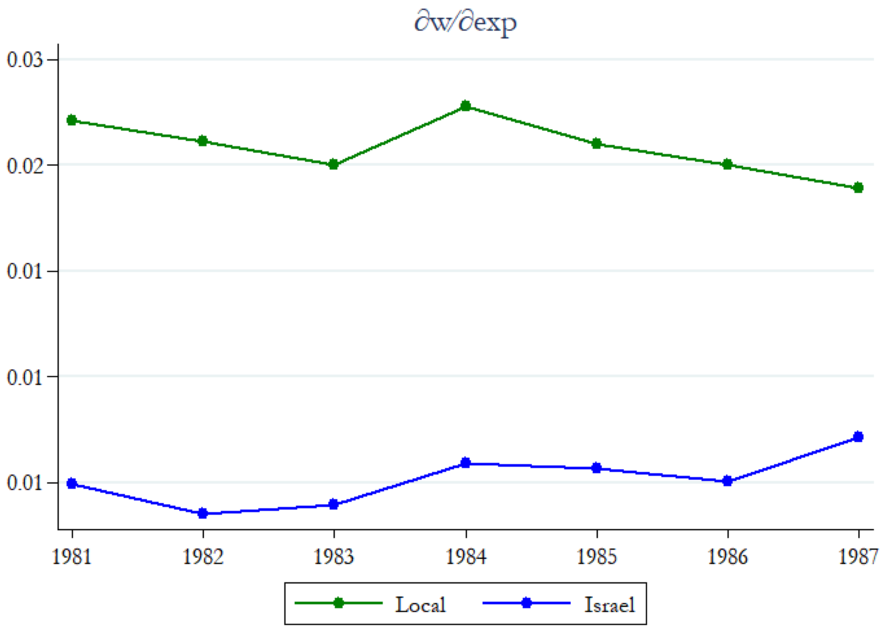
Parameter estimates are taken from Table 3.

11 Figures

Figure 1 Point Estimates of Skills Returns



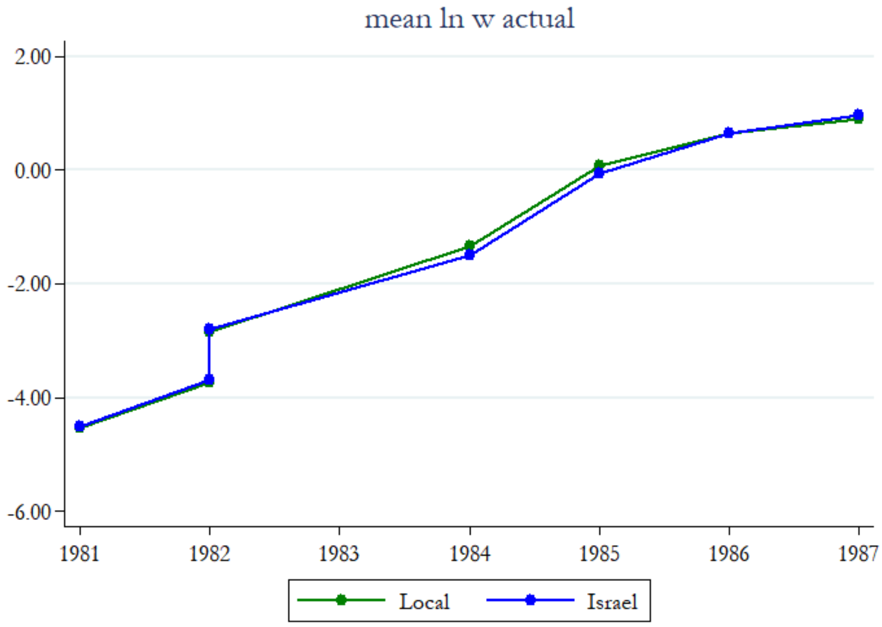
a. Returns on education



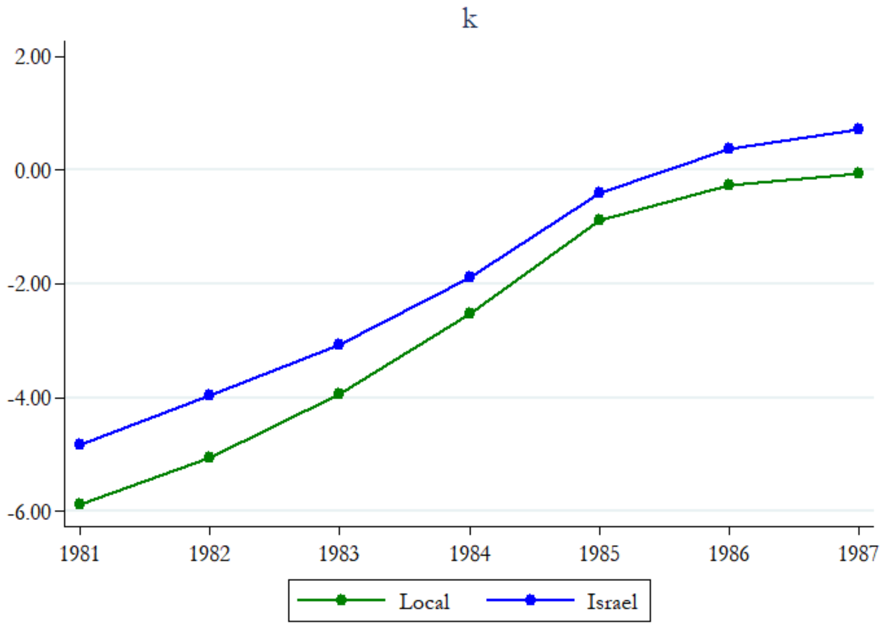
b. $\frac{\partial \ln w}{\partial \text{experience}}$

Notes:
1. Based on Table 3.

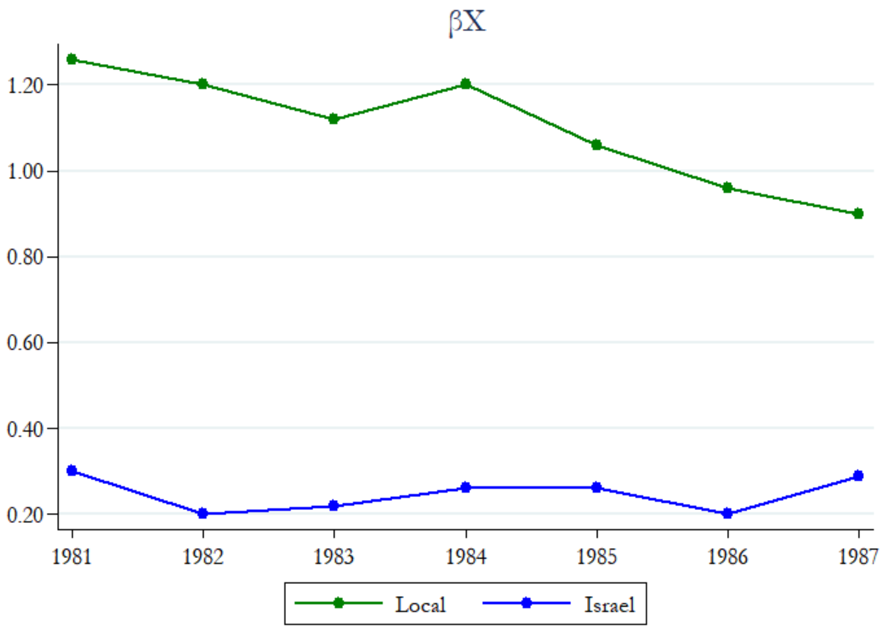
Figure 2: Log Wage Decompositions



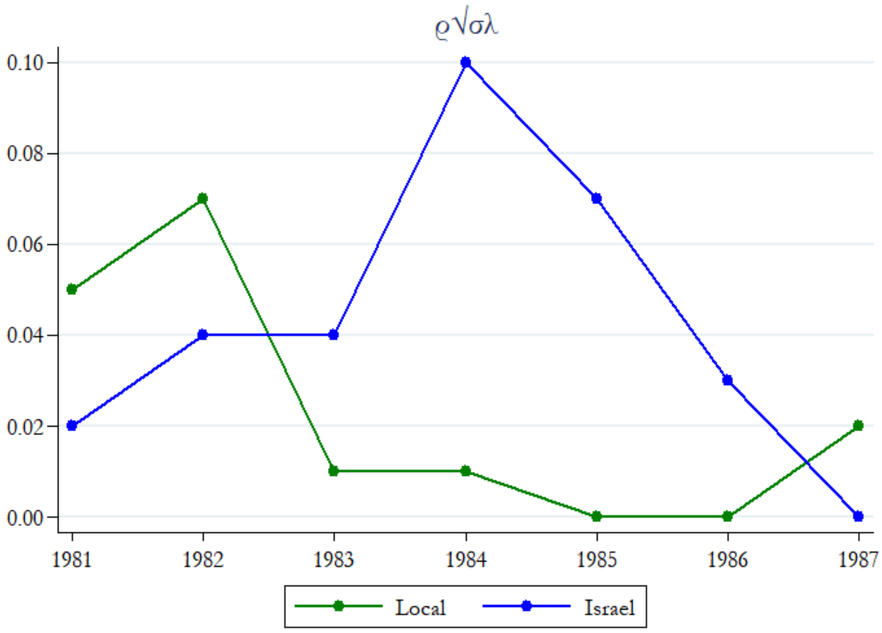
a. **Mean Log Wages**



b. Wage Equation $\widehat{k}$



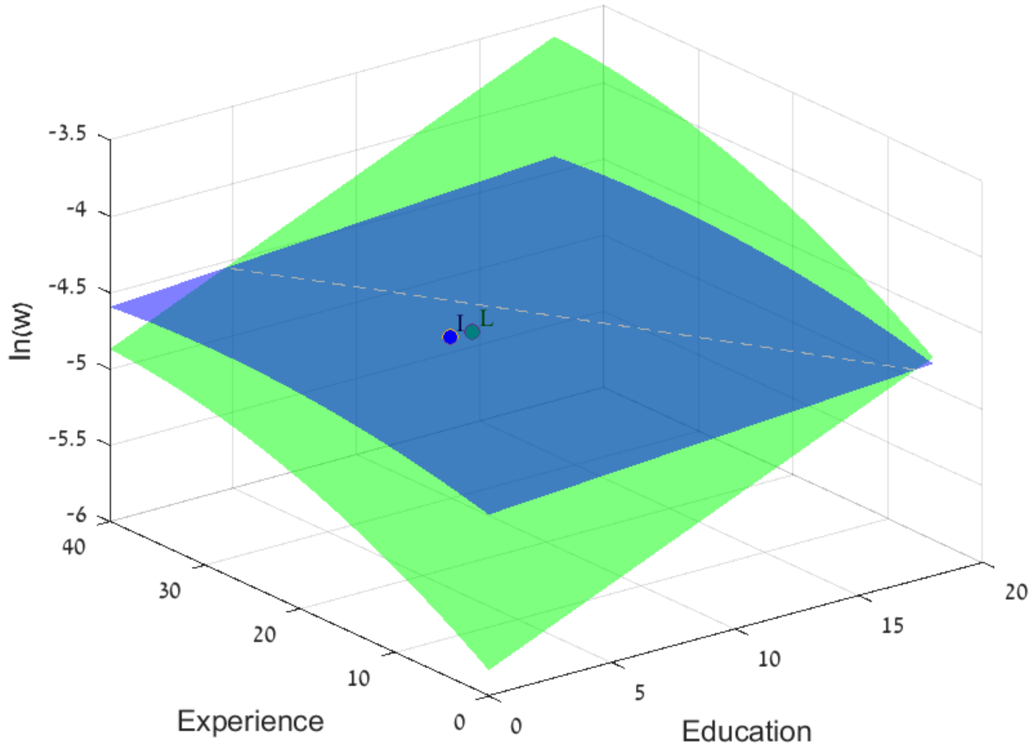
c. $\widehat{\beta}\overline{X}$



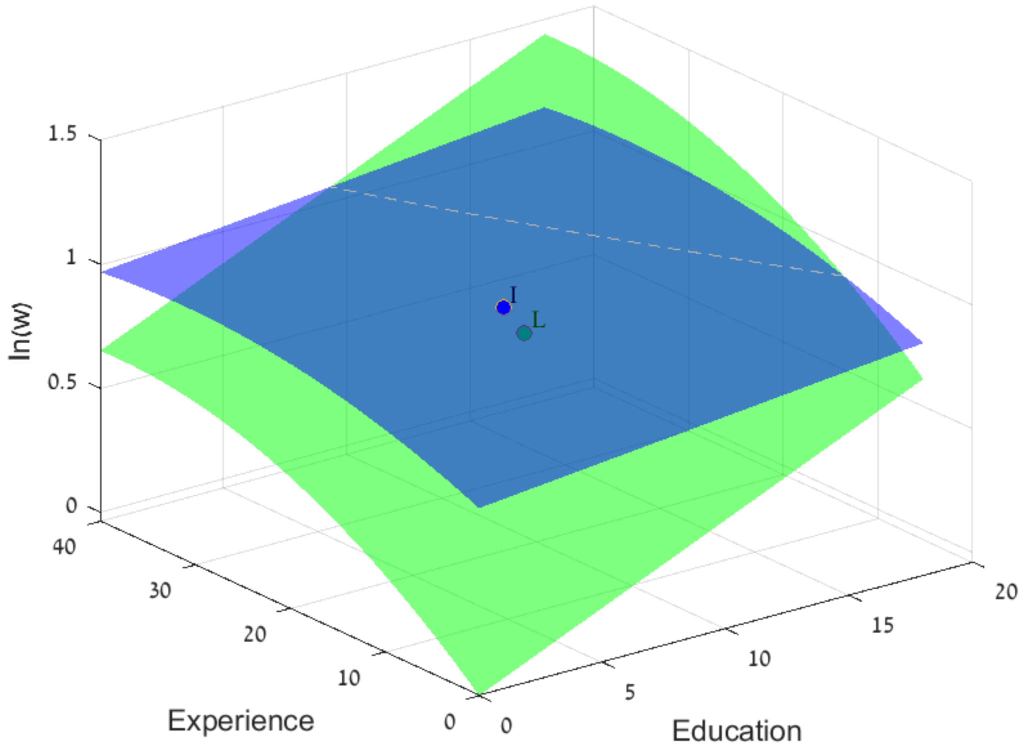
d. Selection term $\widehat{\rho}\widehat{\sqrt{\sigma}}\widehat{\lambda}$

Notes:
1. Based on Table 5.

Figure 3: Wages as a Function of Skills



a. 1981 log wage equation



b. 1987 log wage equation

\

Notes:

a. Graphs depict the equation

$$\ln w_i \mid \left(E\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\widehat{\lambda}_i, Ev_i\right) = C_i + \beta^{educ} educ_i + \beta_1 S_{h,i} + \beta_2 S_{h,i}^2 \tag{29}$$

where $S_{h,i}$ is experience and where

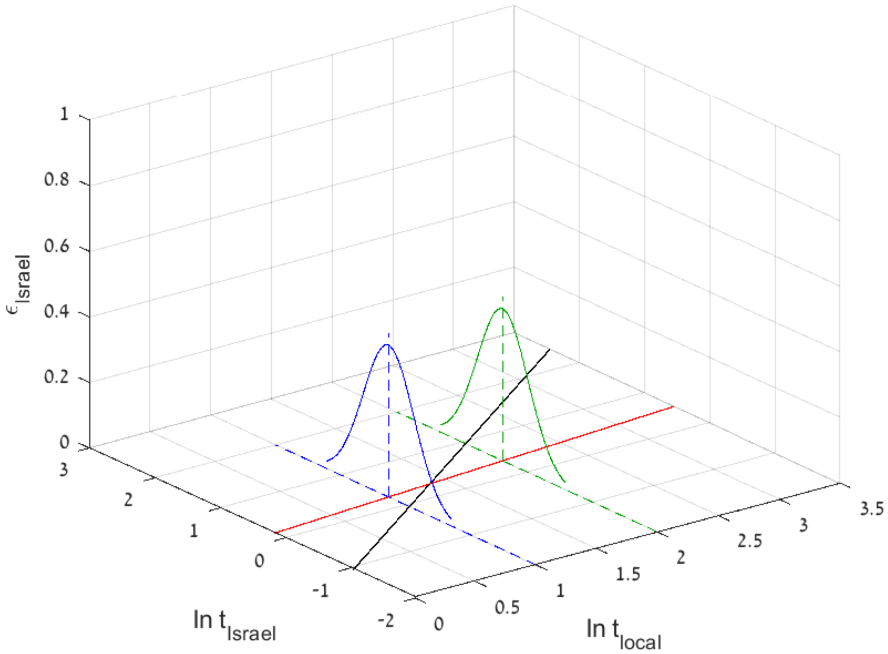
$$C_i = \widehat{k}_i + \beta_{i,0} + E\left(\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\widehat{\lambda}_i\right)$$

The estimates are taken from Table 3.

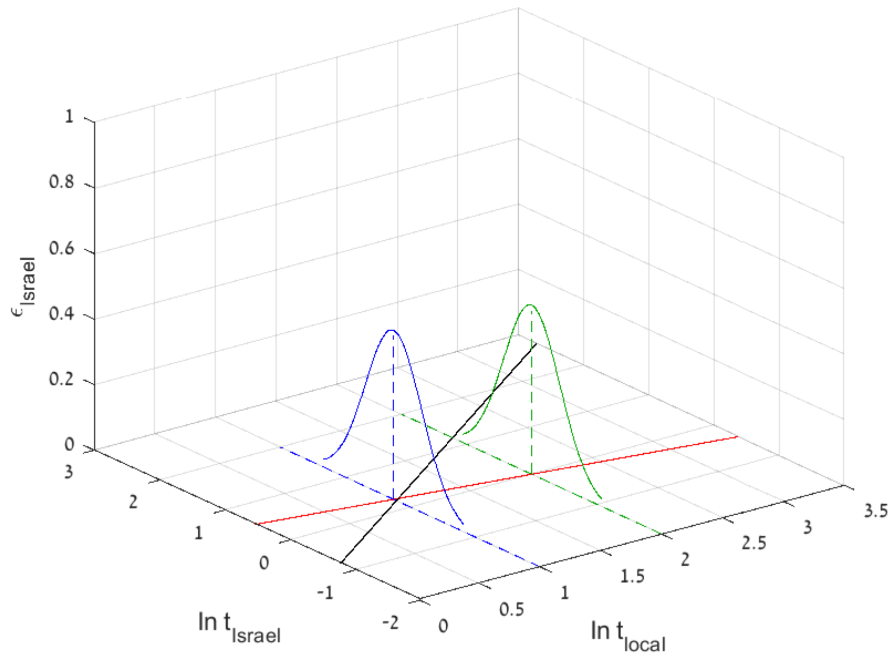
b. Blue marks workers in Israel and green marks local workers.

c. The points L, I mark the values presented in Table 7 for workers in the local and Israeli economy, respectively.

Figure 4: Tasks and Unobserved Skills



a. 1981 estimates



b. 1987 estimates

Notes:

1. Equation (26) 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 1.

12 Appendix A: Derivation of Product Output per Worker and Wages in a Location

I postulate a Cobb Douglas production function for product output in location i :

$$Y_i = K_i^{\alpha}(A_i T_i)^{1-\alpha} \tag{30}$$

This implies:

$$\frac{K_i}{Y_i} = \frac{K_i}{K_i^{\alpha}(A_i T_i)^{1-\alpha}} = \frac{K_i^{1-\alpha}}{(A_i T_i)^{1-\alpha}} \tag{31}$$

$$\left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} = \left(\frac{K_i}{A_i T_i}\right)^{\alpha} \tag{32}$$

Product output per worker is thus given by:

$$\begin{aligned} \frac{Y_i}{L_i} &= \left(\frac{K_i}{L_i}\right)^{\alpha}(A_i \frac{T_i}{L_i})^{1-\alpha} \\ &= \left(\frac{K_i}{A_i T_i}\right)^{\alpha} A_i \frac{T_i}{L_i} \end{aligned} \tag{33}$$

Using the relation $T_i = L_i \widetilde{T}_i$ and equation (32):

$$\frac{Y_i}{L_i} = \left(\frac{K_i}{Y_i}\right)^{\frac{\alpha}{1-\alpha}} A_i \widetilde{T}_i \tag{34}$$

Define:

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

So:

$$\ln \frac{Y_i}{L_i} = \ln z_i + \ln \widetilde{T}_i \tag{36}$$

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

$$\begin{aligned} \ln w_i &= \ln(1-\alpha) + \ln \frac{Y_i}{L_i} \\ &= \ln(1-\alpha) + \ln z_i + \ln \widetilde{T}_i \end{aligned} \tag{37}$$

13 Appendix B: Econometric Methodologies

I use two alternative methods to estimate equations (11) for workers employed locally and employed in Israel as follows.

13.1 Heckman Self-Selection Model

Following Heckman (1979) and Heckman and Sedlacek (1985) I proceed as follows.

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:

$$\ln w_i = \ln \pi_i + \beta_i \mathbf{X} + u_i, \tag{38}$$

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

When estimating equation (38), I take into account sample selection, which is inherent in the model. Thus define the variable z^* :

$$\begin{aligned} z^* &= \ln w_i + \ln(1 - k_i(\mathbf{L})) + \ln(1-\gamma_i) - \ln w_j - \ln(1-k_j(\mathbf{L})) - \ln(1-\gamma_j) \quad (39) \\ &= \ln \pi_i - \ln \pi_j \\ &\quad + \ln(1 - k_i(\mathbf{L})) - \ln(1-k_j(\mathbf{L})) \\ &\quad + \ln(1-\gamma_i) - \ln(1-\gamma_j) \\ &\quad + \beta_i \mathbf{X} - \beta_j \mathbf{X} \\ &\quad + u_i - u_j \end{aligned}$$

and the indicator variable z :

$$\begin{aligned} z &= 1 \; if \; z^* > 0 \\ z &= 0 \text{ otherwise} \end{aligned} \tag{40}$$

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

$$\begin{aligned} \Pr(z &= 1) = \Phi(\ln \frac{\pi_i}{\pi_j} + \ln \frac{(1 - k_i(\mathbf{L}))}{(1-k_j(\mathbf{L}))} + \ln(1-\gamma_i) - \ln(1-\gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \quad (41) \\ \Pr(z &= 0) = 1 - \Phi(\ln \frac{\pi_i}{\pi_j} + \ln \frac{(1 - k_i(\mathbf{L}))}{(1-k_j(\mathbf{L}))} + \ln(1-\gamma_i) - \ln(1-\gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \end{aligned}$$

The observed $\ln w_i$ is given by:

$$\ln w_i \mid (z = 1) = \ln \pi_i + \beta_i \mathbf{X} + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*} \right] \lambda(c_i) + u_i \tag{42}$$

where:

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

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

This may also be written as follows:

$$
\ln w_i \mid (z=1) = \ln \pi_i + \beta_i \mathbf{X} + \rho_i \sqrt{\sigma_{i\,i}} \lambda\,(c_i\,) + u_i \tag{43}
$$

A similar equation holds true for the other location. Note that while the **X** vector appears in both (41) and (43), the **L** vector appears only in the selection equation (41). I estimate the model using Heckman's (1979) two-step consistent estimation procedure. One can interpret the selection bias in (38) 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**:

$$
\frac{\partial E(\ln w_i \mid (z=1))}{\partial x_k} = \beta_{ik} + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}\right] \frac{\partial \lambda}{\partial c_i} \frac{\partial c_i}{\partial x_k} \tag{44}
$$

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.

Identification issues are discussed in the main text, in sub-section 5.2.1.

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 (41). 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[29]I also use indicator variables for the quarters within 1987, which I do not report.

Approximating I get:

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

The selection equations are:

$$\begin{aligned} \Pr(z &= 1) = \Phi(\ln \frac{\pi_i}{\pi_j} + \sum_p \theta_p \cdot l_p^j - \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^j - \sum_n \gamma_n Y_n^i \quad (45) \\ &+ \ln(1-\gamma_i) - \ln(1-\gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \\ \Pr(z &= 0) = 1 - \Phi(\ln \frac{\pi_i}{\pi_j} + \sum_p \theta_p \cdot l_p^j - \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^j - \sum_n \gamma_n Y_n^i \\ &+ \ln(1-\gamma_i) - \ln(1-\gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \end{aligned}$$

The estimated wage equation is the following:

$$\begin{aligned} \ln w_i \quad | \quad & \text{location } i = \ln \pi_i + \beta_{0i} + \beta_{1\,i} educ + \beta_{2\,i} exp + \beta_{3i} exp^2 \\ & + \sum_{m=2}^{4} \gamma_m Q_m + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}\right] \lambda(c_i) + u_i \qquad (46) \end{aligned}$$

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 hourly wages ($\ln w_i$), defined as the nominal monthly wage divided by hours worked. The use of hourly wages is designed to avoid confounding the choice of work place with the choice of work time (hours or days).[30] Education (*educ*) and experience (*exp*) are defined in years.

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

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

[30]The sample includes all wage earners except those with hourly wages below the lowest 1% or above the highest 0.2%. For the deleted 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.

13.2 Semi-Parametric Estimation

I use the methodology proposed by D'Haultfoeuille, Maurel, and Zhang (2018) and D'Haultfoeuille, Maurel, Qiu, and Zhang, (2019) to estimate the model without relying on exclusion restrictions.[31]

The rationale of their methodology is as follows[32]:

> ... in practice, valid instruments are generally difficult to find. Identification at infinity has been proposed in the literature as an alternative solution to the endogenous selection problem, in situations where one is primarily interested in estimating the effects of some covariates on a potential outcome...
>
> D'Haultfoeuille and Maurel (2013) show that identification in the absence of an instrument or a large support covariate is in fact possible. Their key condition is that selection becomes independent of the covariates at infinity, i.e., when the outcome takes arbitrarily large values. The idea behind is that if selection is indeed endogenous, one can expect the effect of the outcome on selection to dominate those of the covariates, for sufficiently large values of the outcome...

The implementation is formally described as follows [33]:

> Specifically, denoting by Y^* and X_1 the outcome and covariates of interest, and by X_{-1} other covariates... we consider the following outcome equation:
>
> $Y^* = X_1'\beta_1 + \varepsilon$
>
> where, for any $\tau \in (0,1)$, the τ-th conditional quantile of ε satisfies $Q_{\varepsilon|X}(\tau|X) = \beta_0(\tau) + X_{-1}'\beta_{-1}(\tau)$.
>
> Denoting by D the selection dummy, the econometrician only observes $(D, Y = DY^*, X)$. In this framework, the effect of interest β_1 is identified from the analysis of D'Haultfoeuille and Maurel (2013)...we extend their result by directly relating β_1 to the upper conditional quantiles of Y. Following this new constructive identification result, we then develop a consistent and asymptotically normal estimator of β_1. We propose an estimator based on extremal quantile regression, that is quantile regression applied to the upper tail of Y...Throughout the paper we focus on the intermediate order case, which corresponds to situations where the quantile index goes to one as the sample size tends to infinity, but at a slower rate than the sample size.

---

[31] Beyond the cited references, see the paper entitled "Estimating Selection Models without Instrument with Stata" by Xavier D'Haultfœuille, Arnaud Maurel, Xiaoyun Qiu, and Yichong Zhang in the Stata Journal, 2020 forthcoming, for the relevant software code.

[32] D'Haultfoeuille, X., A. Maurel, and Y. Zhang (2018 pp.129-130).

[33] D'Haultfoeuille, X., A. Maurel, and Y. Zhang (2018 p.130).

The value added of this method is explained as follows:

> Unlike prior estimation methods for sample selection models, we propose a distribution-free estimator that does not require an instrument for selection nor a large support regressor. Besides and importantly, we do not restrict the selection process, apart from the independence at infinity condition mentioned above. In the context of standard selection models, this condition translates into a restriction on the dependence between the error terms of the outcome and selection equation, which is mild provided that selection is indeed endogenous. The structure of the outcome equation, which generalizes the standard location shift model by allowing for heterogeneous effects of the covariates X_{-1} on different parts of the distribution, also plays an important role for identification...
>
> Importantly, these assumptions are testable, since they imply that for large quantile indices, the estimators of β_1 obtained using different quantile indices are close.

Using this methodology the current paper estimates the following equation, estimated separately for each location:

$$\widetilde{\ln w} = \beta_1 \widetilde{educ} + \beta_2 \widetilde{\exp} + \beta_3 \widetilde{\exp}^2 + u$$

where tilde denoted de-meaned variables, taking into account quarterly dummies.