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

Eran Yashiv†
Tel Aviv University, CfM (LSE), and CEPR

March 10, 2020

Abstract

The phenomenon of workers moving from a poor to a rich economy is high on the political agenda. When a worker moves to a richer economy, what is gained by the move?

The empirical challenge in giving an answer stems from the difficulty to disentangle income differences from many other determinants. Estimates are potentially biased due to substantial misspecification of the model, when omitting relevant determinants.

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 other relevant factors. It explicitly addresses the question of what workers newly experience in the richer economy (higher productivity), what is taken from the poorer economy (human capital), and their choices in moving (self-selection). Importantly, it encompasses the constraints placed on workers in terms of the human capital skills demanded.

The findings show that income differences affecting worker choice are made up of contradictory elements. Consistently with findings in the development accounting literature, 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. But lower job task values operate to lower wages for movers, who are offered manual tasks in the rich economy. The latter loss offsets the former gain. The paper emphasizes the idea that tasks are tied to locations. Workers choose a location-wage-task 'pack,' with movers getting low rewards to the skills bundled in their job tasks.

*I am grateful to Richard Blundell, Francesco Caselli, Chad Jones, John Kennan, and David Lagakos for useful conversations, to Arnaud Maurel for econometric advice, to seminar participants at UCL and Tel Aviv for feedback and suggestions, and to Elad Demalach, Oriel Nofekh, and Alon Rieger for research assistance. Any errors are my own.

†yashiv@tauex.tau.ac.il; e.yashiv@lse.ac.uk

Key Words: development accounting, movers and stayers, rich and poor economies, pure income effects, job tasks, TFP differentials, human capital differences, self-selection, skill bundle.
JEL Codes: E24, J24, J31, O15.

Non Technical Summary

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 or migration across countries. It is a salient issue, with such migration flows very high on the political agenda in many rich countries. When a worker moves to an economy richer than the home economy, what is gained by the move? This question is evidently important for the moving workers themselves, as well as for evaluating both economic gains and political implications. This question is at the focus of the paper.

There is a view whereby such gains are very large. For example, Kennan (2013) estimates a gain in net income of 125% within a model of migrants from poor countries to rich ones. It is not straightforward, however, to answer the question of the gain from the move to a rich country. The difficulty is related to the need to disentangle the effects of income differences on the movers' decision 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, and more. Many estimates in the literature are potentially biased due to substantial mis-specification of the model, when omitting relevant determinants.

This paper studies a unique case that allows to isolate the pure effects of income differences. A key rationale underlying the analysis is akin to the one explored and debated in the development accounting literature – the distinction between factors which are external to the worker, such as technology, capital, and institutions, and factors embodied in the worker, such as skills and abilities. Hence, when estimating wage equations so as to infer the gains from a poor to rich economy move, it explicitly address the question of what workers newly experience in the richer economy (say, higher technology), what is taken from the poorer economy (human capital), and their choices in moving (self-selection). Importantly, it takes into account the fact that movers and stayers are typically constrained in terms of the jobs offered and the skills required.

The unique data set used consists of repeated cross-sections of a Labor Force Survey of Palestinian workers who were working in Israel. The survey sampled both movers and stayers within a unified setting. 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 determinants of mobility 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 be estimated without bias. Indeed "moving" was minimal and the main feature of the data to be exploited here is that workers belonging to a poor economy worked in a rich one.

It should be emphasized that there are two intertwined issues here: one is the estimation of effects of factors external to the worker (technology, capital) as distinct from factors embodied in the workers (skills); the second is that such estimates require appropriate data that eschew the bias inherent in confounding determinants of mobility.

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. Thus, the paper connects with work that breaks down the cross-country differences into the share of technology and capital and the share of human capital. The distinction between the two is key.

Second, it explicitly recognizes that workers face job tasks requirements and particular rewards for their skills in performing these tasks. Here this paper connects with Autor and Handel (2013), who estimate a similar self-selection model with U.S. job and wage data. They note the issue of skill bundling within tasks. The bundling in the current paper is in terms of location-task-skills. Workers are demanded for a particular task,which requires a bundle of skills and rewards it in a specific way in a particular location.

I use two alternative estimation methodologies to examine wage regressions of movers and stayers. I analyze the findings across the two economies both in terms of the mean wage differential and in terms of the distributions involved.

My findings offer a new take on 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 overturning it. Self-selection on unobservables, which is positive, 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. These findings also imply that the self-selection of movers in terms of skills is not the unique major element here. The productivity differences involved are no less important. Knowing the patterns of self-selection does not suffice to understand the poor to rich move.

The contribution of this paper consists of the following: the literature often looks at the move from poor to rich economies (i) without disentangling the income differences motive from a plethora of other motives; and (ii) anticipates a big productivity gain due to the rich economy having higher TFP and capital. This paper shows that with respect to point (i), there is potential for substantial misspecification and bias, while the unique data set used here eschews such bias. With respect to point (ii), the paper emphasizes the idea that tasks are tied to locations, and so workers choose a location-wage-task 'pack' that determines rewards to the skills bundled in the task. The bundling constrains the human capital returns for movers and generates a big offset to the productivity gain.

The analysis of this paper is relevant for many cases of foreign minorities in advanced economies. Workers belonging to such minorities are demanded to perform low-skill tasks, as is the case here. 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.

Moving from a Poor Economy to a Rich One:
The Contradictory Roles of Technology 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.[2] It is a salient issue, with such migration flows very high on the political agenda in many rich countries. When a worker moves to an economy richer than the home economy, what is gained by the move? This question is evidently important for the moving workers themselves, as well as for evaluating both economic gains and political implications. This question is at the focus of the paper.

There is a view whereby such gains are very large. For example, Kennan (2013) estimates a gain in net income of 125% within a model of migrants from poor countries to rich ones. This view may find reinforcement in some findings of the development accounting literature. This literature uses GDP per capita and wage data to compute cross-country differences, which are found to be very large, especially between poor and rich economies; see, for example, Jones (2016).

It is not straightforward, however, to answer the question of the gain from the move to a rich country. The difficulty is related to the need to disentangle the effects of income differences on the movers' decision 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, and more. Many estimates in the literature are potentially biased due to substantial mis-specification of the model, when omitting relevant determinants.

This paper studies a unique case that allows to isolate the pure effects of income differences. A key rationale underlying the analysis is akin to the one explored and debated in the development accounting literature – the distinction between factors which are external to the worker, such as TFP,

[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]Consider two measures.

(i) The permanent immigration *flows* into the G7 countries in 2016 was 3.4 million (OECD (2020)) out of roughly 7.5 million immigrants globally, i.e., 45%.

(ii) In 2019 out of an estimated *stock* of 130.2 million migrants worldwide, 51.9 million originated from less developed regions by UN classification (40%) and 46. 5 million originated from non-high income countries by World Bank classification (36%). Source: UN (2019).

capital, and institutions, and factors embodied in the worker, such as skills and abilities. Hence, when estimating wage equations so as to infer the gains from a poor to rich economy move, it explicitly address the question of what workers newly experience in the richer economy (say, higher TFP), what is taken from the poorer economy (human capital), and their choices in moving (self-selection). Importantly, it takes into account the fact that movers and stayers are typically constrained in terms of the jobs offered and the skills required.

The unique data set used consists of repeated cross-sections of a Labor Force Survey of Palestinian workers who were working in Israel. The survey sampled both movers and stayers within a unified setting.[3] 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 determinants of mobility 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 be estimated without bias. Indeed "moving" was minimal and the main feature of the data to be exploited here is that workers belonging to a poor economy worked in a rich one.

It should be emphasized that there are two intertwined issues here: one is the estimation of effects of factors external to the worker (TFP, capital) as distinct from factors embodied in the workers (skills); the second is that such estimates require appropriate data that eschew the bias inherent in confounding determinants of mobility.

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. Thus, the paper connects with the work of Hendricks and Schoellman (2018, 2019), reviewed below, that breaks down the cross-country differences into the share of technology and capital and the share of human capital. The distinction between the two is key.

Second, it explicitly recognizes that workers face job tasks requirements and particular rewards for their skills in performing these tasks. Here this paper connects with Autor and Handel (2013), who estimate a similar self-selection model with U.S. job and wage data. They note the issue of skill bundling within tasks. The bundling in the current paper is in terms of location-task-skills. Workers are demanded for a particular task,which requires a bundle of skills and rewards it in a specific way in a particular location.

[3]This data set was processed and used by Angrist (1995, 1996) to study other issues.

I use two alternative estimation methodologies to examine wage regressions of movers and stayers. I analyze the findings across the two economies both in terms of the mean wage differential and in terms of the distributions involved.

My findings offer a new take on 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 technology 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 overturning it. Self-selection on unobservables, which is positive, 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. These findings also imply that the self-selection of movers in terms of skills is not the unique major element here. The productivity differences involved are no less important. Knowing the patterns of self-selection does not suffice to understand the poor to rich move.

The contribution of this paper consists of the following: the literature often looks at the move from poor to rich economies (i) without disentangling the income differences motive from a plethora of other motives; and (ii) anticipates a big productivity gain due to the rich economy having higher technology and capital. This paper shows that with respect to point (i), there is potential for substantial misspecification and bias, while the unique data set used here eschews such bias. With respect to point (ii), the paper emphasizes the idea that tasks are tied to locations, and so workers choose a location-wage-task 'pack' that determines rewards to the skills bundled in the task. The bundling constrains the human capital returns for movers and generates a big offset to the productivity gain.

The analysis of this paper is relevant for many cases of foreign minorities in advanced economies. Workers belonging to such minorities are demanded to perform low-skill tasks, as is the case here. 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.

The paper proceeds as follows. Section 2 offers the background and context in the literature. It highlights insights from the development accounting literature and discusses the aspects of the task approach in labor markets that are pertinent here. Section 3 presents the model. It elaborates on the role of skills and tasks in the model and offers a connection with the analysis of the development accounting literature. Section 4 presents the Palestinian labor market and its key features, so as to justify the use of certain elements of the model (presented in Section 3). It then discusses the data set and presents summary statistics. Section 5 presents the two econometric methodologies and the results, with some further elaboration in the appendix. Section 6 analyzes the components of the movers-stayers mean wage differentials and their significance. Section 7 examines the skills and wage distributions across locations and their implications for the moving decision. Section 8 discusses the results in terms of their implications for the development accounting literature and in terms of their applicability to other cases of moving from poor to rich economies. Section 9 concludes.

2 Literature

This paper relates to two strands of literature. The first is the development accounting literature, discussed below in sub-section 2.1, which studies cross-country income differences. The second is the task-based approach to the labor market, discussed below 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. This issue is important for the current paper in the sense that it decomposes movers wage gains into a TFP cum capital part and a human capital related part. That said, there are, however, important differences between the DA approach and the approach here, to be discussed in sub-section 8.1 below.

Caselli (2005, 2016) 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.

In terms of the breakdown into components, 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 this debate on the relative size of TFP vs human capital shares. Their work presents evidence from the experiences of immigrants to the U.S. The underlying logic is that immigrants enter the U.S with the human capital they acquired in their birth country and work with U.S physical capital and TFP. Their wage gains compared to GDP per capita differences allows to separate the human capital factor from these country-specific factors. 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) extend this 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 earlier findings. The share of output per capita 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 in Jones (2019).

There are papers in the migration literature, focusing on migration from poor to rich economies, which relate to similar questions. In a prominent contribution in this context, 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%.

The common thread of these various studies, and the issue that is relevant for the current paper, is the distinction between the environment in which the worker operates (technology, capital, and the related institutions) and what is embodied in the worker (skills and abilities). In this paper I discuss my findings of movers' wage gains in terms of the distinct components of human capital and of TFP and physical capital. I suggest a mechanism, to account for the results, which has not been evaluated by the afore-cited 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 workers even if holding several jobs can perform tasks only in one job at a time. The indivisible bundling of tasks within jobs implies that the productivity of particular task inputs will not necessarily be equated across jobs. Using U.S. job and task data, they test the model's predictions for this relationship, finding empirical support for the model. Within this line of inquiry, Gathmann and Schönberg (2010) develop and test the empirical implications of a setting in which workers differ in their productivity across tasks and task returns differ across occupations. They study the implications for the evolution of job changes and wage growth.

In the model here, I, too, 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 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 discuss more in depth the role of tasks and skills in this model; finally, in sub-section 3.3, 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. Packages of skills cannot be unbundled and different skills are used in different 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. The next section provides more details.

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

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 $\gamma_i = \gamma \in [0,1)$ 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. 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:

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

where

$$\begin{aligned} \mu_i &= \beta_{i,0} + \sum_h \beta_{h,i} S_h \\ \mu_j &= \beta_{j,0} + \sum_h \beta_{h,j} S_h \end{aligned}$$

With these functional specifications, the following holds true:[4]

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

where

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

and $\Phi(\cdot)$ 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

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

costs decline, i.e. as $\ln(1-\gamma_i) - \ln(1-\gamma_j) + \ln \frac{[1-k_i(\mathbf{L})]}{[1-k_j(\mathbf{L})]}$ 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 The Role of Tasks and Skills

The role of tasks and skills here merits some elaboration. Heckman and Sedlacek (1995), point out that skills cannot be unbundled and that different skills are used in different tasks. Hence the wage equation, such as (11) above, is not a conventional hedonic function, whereby implicit market prices out each component of the skill vector. In the model here tasks are priced, not the components of the skill vector directly.

Autor and Handel (2013) discuss this issue further. They note (see their pages S62, S64, and S66) that job tasks are not fixed worker attributes. Rather, workers modify their task inputs by self-selecting into jobs according to comparative advantage and reallocate their labor input among tasks when the market value of tasks changes. Tasks are not durable investment goods that earn a well-defined market rate of return. The tasks that a worker performs on the job are an application of the workers skills to a given set of activities, and workers can modify task inputs as job requirements change. Hence there will not be a one-to-one mapping between a worker's stock of human capital and the job tasks performed. Tasks are a high-dimensional bundle of activities, the elements of which must be performed jointly to produce output. Core job tasks cannot be unbundled and each worker in the job must perform them. The equilibrium of the model ensures that workers are employed in the location that has the highest reward to their bundle of tasks. But this does not imply that they receive the maximum market reward to each element in their task bundle or that each element is equally valuable in all locations.

The bundling in the current paper is in terms of location-task-skills. Workers are demanded to perform a particular task, which requires a bundle of skills and rewards it in a specific way, in the particular location. They cannot choose to perform this task in another location, or perform another task in a given location.

The significance of this set up will become clear below. As will be seen, the returns to skills are not the same across the two locations. This differential plays a major role in conjunction with the differential in productivity (stemming from TFP and capital differences) across locations.

3.3 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 at the outset a crucial distinction with respect to this literature. In the current paper, $\ln w_i$ always refers to a wage of a Palestinian worker, not an Israeli worker, and the index i refers to the location – Israel or the local economy. Hence wage gains are going to be empirically examined across locations 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:

$$\begin{aligned} z_i &\equiv K_i^{\alpha} A_i^{1-\alpha} T_i^{-\alpha} \\ &= \left(\frac{K_i}{T_i}\right)^{\alpha} A_i^{1-\alpha} \end{aligned} \tag{14}$$

where z_i is a function of the aggregate variables K, T and A in location i.

In logs:

$$\ln z_i = \alpha \ln \frac{K_i}{T_i} + (1-\alpha)\ln A_i$$

Given that

$$Y_i = z_i T_i$$

Using (4), one gets that the task price π_i, equals the productivity measure, z_i, in the location:

$$\pi_i = \frac{\partial F^i}{\partial T_i} = z_i \tag{15}$$

Using equation (7) this means:

$$\begin{aligned} \ln w_i(\mathbf{S}) &= \ln \pi_i + \ln t_i(\mathbf{S}) \\ &= \ln z_i + \ln t_i(\mathbf{S}) \end{aligned} \tag{16}$$

Estimation of the log wage equation will provide estimates of z_i, facilitating comparisons with the findings of the development accounting literature. Note, though, 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.

Workers can gain by a move to a richer economy with a higher level of z_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. 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 Palestinian Labor Market and the Data

I describe the features of the Palestinian labor market. In particular, the changes over time in the flows of Palestinians into Israel, and in restrictions imposed or lack of them, are key in determining the choice of the time period for the data sample used. In what follows I draw on Semyonov and Lewin-Epstein (1987), Angrist (1995, 1996), Arnon, Luski, Spivak, and Weinblatt (1997, in particular Chapter 3), and Bartram (1998).[5]

Palestinian workers in Israel. 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. The share of salaried employees employed in Israel at 22% in 1970, climbed to around 50% in the first half of the 1970s, and then fluctuated around that rate and up to 65%, starting to fall off in the late 1980s. Hence, a key employment decision of the Palestinian male worker was the choice of employment location – Israel or the local economy. Men 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 subject to restrictions, much more volatile and, generally, on a declining trend.[6]

[5] For an analysis of the Israeli labor market, see Yashiv (2000).

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

Patterns of Commuting and Employment. In the first few years following 1967, the flow of Palestinian workers into Israel was regulated through the issue of work permits and through centralized payment arrangements. But the market gradually became unrestricted and de-regulated by the end of the 1970s, when employment in Israel increased considerably. Palestinian workers were commuting to work, travelling 30-90 minutes to work daily or weekly. If staying in Israel more than a day, they were lodged in low-quality housing, close to the site of employment. Travel and housing were provided by the employers or by middlemen, and their costs were deducted from wages.[7]

Tying The Model Formulations to Empirical Findings. Angrist (1996) estimates a short run demand function for Palestinian workers. He uses a competitive model (see his pages 437-439) and implements it empirically using data which is taken from the same survey but relates to a later time period (relative to the current paper). The χ^2 goodness of fit statistics suggest a good fit (see his Table 4 on page 447). The model here accords with this empirical work in the sense that there was a well-behaved demand function for Palestinian workers in Israel in a competitive setting.

The Rich and Poor Economies Context. 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 currency 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. 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. As mentioned above, this situation was different prior to the sample period (more regulated) and changed radically after it, starting in the first uprising in December 1987. I abstain, purposefully, from using post-1987 data, which did feature extensive and time-varying restrictions on the employment of Palestinian workers in Israel. The model below relates to two groups – movers and stayers; there was no other major location decision and hence no third group.

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, 1996).[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. Cur-

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

[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

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

rent 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 above.

Table 1 presents sample statistics.

Table 1

The table shows that, for most, but not all, years, local workers (stayers) earned slightly lower wages.[12]Throughout the sample years stayers 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 Bank and Gaza as the other location. In what follows I present the econometric methodologies (5.1) and the results (5.2). At the end of the section I discuss the uniqueness of this data set in terms of eschewing misspecification and potential bias (5.3).

5.1 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 the Appendix; the following is a short summary.

[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.

5.1.1 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:

$$k_i(\mathbf{L}) = \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^i \tag{17}$$

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 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.[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

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

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

with the choice of work time (hours or days).[15] Education (*educ*) and experience (*exp*) are defined in years. The first specification reported below features only the geographical exclusion restrictions. The second specification includes in the set of exclusion restrictions the three variables discussed above. The third specification uses OLS to test for the effect of selection correction, running only the wage equation.

5.1.2 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 quantile regression applied to the upper tail of the outcome variable.[16] The Appendix provides a formal definition.

5.2 Results

Table 2 reports the full results of the Heckman methodology[17] using the two alternative sets of exclusion restrictions, and using OLS, for the TLFS cross-section in 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 slight differences in the estimates of the intercept in Israel employment. The emerging picture across columns 1 and 2 is the same, but column 1 has higher point estimates for the returns to skills. Overall, the differences in point estimates across specifications are not substantial. In what follows, I use as the benchmark the specification of column 1, i.e., the one with the smaller exclusion restrictions set.

[15] The sample truncates the bottom 1% and the top 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).

[16] See Angrist and Pischke (Chapter 7.1, 2009).

[17] I include estimates of the implied second moments and the Wald test (using χ^2 test statistics, with p-values in parentheses).

Table 3 and Figure 1 present the results for the seven repeated cross-sections in the years 1981 to 1987, using this specification.

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 implementing the semi-parametric methodology discussed in sub-section 5.1.2 above. It presents the skill premia estimates,[18] and repeats the results of the Heckman specification (of Table 3 above), for all cross-sections in the years 1981-1987.

Table 4

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.

5.3 Data Uniqueness and Issues of Misspecification Bias

The current model, given the unique data features discussed above, is not subject to potential mis-specification, which is prevalent in many other cases. This is so as generally there may be other determinants, beyond wage differences net of costs, affecting the moving decision. The set up of the current paper *precludes* this possibility. In what follows I show what a model with these other variables entails and the ensuing mis-specification when these determinants are not taken into account. As shown below, this is *not* just a case of omitted variables bias in the wage equation.

[18]This methodology does not facilitate the estimation of the intercept.

5.3.1 Determinants Affecting the Move to a Rich Economy

The analysis in Dao, Docquier, Parsons, and Peri (2018) presents variables that potentially drive the moving decision. In the current case they do not play a role and hence their omission is not problematic, as explained in the following discussion.

Geographical distance. The distance to be travelled is an obvious determinant, affecting costs, including possibly socio-psychological costs. In the current case this was a commute and the distance was travelled, usually daily or 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 bring and 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 and there was no host economy network.

Credit constraints. Credit constraints 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.

5.3.2 Potential Mis-Specification

To understand the potential mis-specification here, the following is a brief version of a generalized Roy model, incorporating the determinants discussed above, implemented to the current setting.

Following the formulation of D'Haultfoeuille and Maurel (2013) of such a Roy model, the non-wage component of the location decision is allowed to vary across individuals and is given by:

$$G_i(X) = \mu_i(X) + U_i \tag{18}$$

whereby $\mu_i(X)$ is the deterministic part, and $U_i \sim N(0, \sigma_U^2)$. X is a vector of variables, and U_i is a distribution, both reflecting the afore-listed variables. Note that $-G_i(X)$ can be interpreted as a cost of moving to location i. It is the $G_i(X)$ function, which captures the effects of the variables discussed in the preceding sub-section above.

Denote by w_i the potential wages in location i and by η_i location specific productivity terms, and so:

$$E(w_i \mid X, \eta_i) = \psi_i(X) + \eta_i \tag{19}$$

Assuming

$$\eta_i \sim N(m_i, \sigma_{\eta_i}^2) \tag{20}$$

Importantly the functions μ_i and ψ_i are not the same and η_i reflects productivity and not non-wage factors of the kind discussed above and captured by U_i. Essentially, $m_i = \pi_i$ in the current model.

Unlike the model presented above, choice in this case is based not only on income maximization. Rather, each worker chooses the location, which yields the highest utility in location i, given by

$$\bar{U}_i = \psi_i(X) + \eta_i + G_i(X) \tag{21}$$

The point is that the current paper posits $G_i(X) = 0$ in line with the data but in many empirical cases this does not hold true.

Heckman and Sedlacek (1985, Appendix B) analytically derive, in their equations B2 and B3,[19] the density of wages in each location, w_i, conditional on the choice $\bar{U}_i > \bar{U}_j$.These conditional wage densities are functions of trivariate normal integrals, which themselves are functions of (inter-alia) the non-wage component $G_i(X)$.Within this latter component, $\mu_i(X)$ and U_i, with its variance σ_U^2, play a role. Thus, the potential mis-specification arises whenever $\mu_i(X) \neq 0$ or $\sigma_U^2 \neq 0$ or both, as is very likely to be the case in numerous applications.

Note, then, that this is not just a case of omitted variables bias in the wage equation. The optimal location selection, based on equation (21), is mis-specified, and, as the object of interest are wages conditional on selection, any estimation of wages is mis-specified. One needs a data set of the kind used in this paper to avoid this state of affairs, or a very rich data set which can allow for the identification of $G_i(X)$.

[19]Assuming particular functional forms for $\psi_i(X)$ and $\mu_i(X)$.

6 Components of Mean 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.[20]

6.1 Decomposition of the Mean Wage Differential

In Table 5 and Figure 2, I quantify the relative role played by the different elements of the model, in terms of means – 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 (see Heckman (1979)):

$$\begin{aligned}
\overline{\ln w_{local}} \quad &| \quad (w_{local} > w_{Israel}) = \widehat{k}_{local} + \widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} + \left(\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\right)\overline{\widehat{\lambda_{local}}} \quad (22)\\
\overline{\ln w_{Israel}} \quad &| \quad (w_{Israel} > w_{local}) = \widehat{k}_{Israel} + \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel} + \left(\widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\right)\overline{\widehat{\lambda_{Israel}}}
\end{aligned}$$

where $\overline{\ln w_i}$ is the mean log hourly wage (conditional on selection) 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 $\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\overline{\widehat{\lambda_i}}$ are the estimates of the second moments $(\widehat{\rho}_i\widehat{\sqrt{\sigma_{ii}}})$ times the average of the estimated inverse of Mills' ratio $(\overline{\widehat{\lambda_i}})$. The table and figure pertain to the repeated cross-sections in the period 1981-1987, using the Heckman methodology.

Table 5 also 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.

$$\begin{aligned}
\overline{\ln w_{local}} \quad &| \quad (w_{local} > w_{Israel}) - \overline{\ln w_{Israel}} \mid (w_{Israel} > w_{local}) \quad (23)\\
&= \quad \widehat{k}_{local} - \widehat{k}_{Israel}\\
&\quad\; + \overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel}) + \widehat{\boldsymbol{\beta}}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})\\
&\quad\; + \left(\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\right)\overline{\widehat{\lambda_{local}}} - \left(\widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\right)\overline{\widehat{\lambda_{Israel}}}
\end{aligned}$$

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

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 $\left(\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\right)\overline{\widehat{\lambda_{local}}} - \left(\widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\right)\overline{\widehat{\lambda_{Israel}}})$.

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

The mean wage differential in the data. The data show 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[21]are higher in the local economy, as seen in Tables 3 and 4 and in Figure 2.[22]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 (23) 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 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, presented below.

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 econ-

[21]The table makes use of point estimates. The linear-quadratic experience premia *profile* in the local economy, shown in Figure 1 above, lies well above that of Israel.

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

omy. 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. Importantly, 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 U.S. task and job data. This last point is key, as will be shown in the interpretation of the results against the background of the findings of the development accounting literature.

Selection on Unobservables. The last term in equation (23), $\left(\widehat{\rho}_{local}\sqrt{\widehat{\sigma_{local}}}\right)\overline{\widehat{\lambda_{local}}} - \left(\widehat{\rho_{Israel}}\sqrt{\widehat{\sigma_{Israel}}}\right)\overline{\widehat{\lambda_{Israel}}}$, ranges between -0.09 and $+0.03$. It thus contributes relatively little to the explanation of the wage differential across location. The next sub-section goes into detail about the type of selection involved here.

Accounting for the Wage Differential. The afore-going discussion portrays 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 in four sample years 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 as follows:

$$\begin{aligned}
& E\left(\ln w_i \mid \ln w_i + \ln\left[1-k_i(\mathbf{L})\right] + \ln[1-\gamma_i] > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right]\right) + \ln[1-\gamma_j] \\
= & \ln \pi_i + \mu_i + \frac{\sigma_{ii}-\sigma_{ij}}{\sigma^*}\lambda(c_i) \qquad (24) \\
& var\left(\ln w_i \mid \ln w_i + \ln\left[1-k_i(\mathbf{L})\right] + \ln[1-\gamma_i] > \ln w_j + \ln\left[1-k_j(\mathbf{L})\right]\right) + \ln[1-\gamma_j] \\
= & \sigma_{ii}\left\{ \begin{array}{c} \rho_i^2[1-c_i\lambda(c_i)-\lambda^2(c_i)] \\ +(1-\rho_i^2) \end{array} \right\} \qquad (25)
\end{aligned}$$

where the inverse Mills ratio is $\lambda(c_i) = \frac{\phi(c_i)}{\Phi(c_i)}$,and $\phi(\cdot)$, $\Phi(\cdot)$ are the density and CDF of a standard normal variable, respectively.

It is possible to classify the selection outcomes in terms of the relations between the elements of Σ: σ_{ii}, σ_{jj} and σ_{ij} or alternatively between $\frac{\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}}}$ and $\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$.[23] Assuming, without loss of generality, that $\sigma_{jj} \geq \sigma_{ii}$, the different outcomes depend on the relation between the ratio of the standard deviation in each location $\frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$ and the correlation between the two locational distributions ρ_{ij}. Three cases are possible:[24](i) positive correlation between the two locations, which is relatively high, i.e., $\rho_{ij} \geq \frac{\sqrt{\sigma_{ii}}}{\sqrt{\sigma_{jj}}}$. In this case, selection is positive in location j and negative in i; (ii) negative correlation between the two locations i.e., $\rho_{ij} < 0$. This is a case of positive selection in the two countries with absolute advantage – each location tends to be filled with the workers that perform best in that 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 due to comparative advantage.

Tables 2 and 3 above report estimates of the unobserved skills variance-co-variance matrix ($\sum$). The results of estimation indicate that (i) the correlation $\rho_{israel,local}$ is lower than the ratio of standard deviations $\frac{\sqrt{\sigma_{israel}}}{\sqrt{\sigma_{local}}}$; moreover, in five out of the seven years it is negative; (ii) the variance in local

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

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

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

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 on tasks. 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. They are consistent with the findings of Autor and Handel (2013) on bivariate relationships between returns on abstract, analytical and on manual skills, 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.

7 Wage and Skills Distributions

Thus far the analysis has pertained to means. What lessons may be drawn from the results of estimation, reported in Table 3, when one considers the entire distributions of wages and of skills? This section undertakes the analysis in these terms, relating to the log wage equation in location i, conditional on selection and its components (following equation (4a') in Heckman (1979)):

$$\ln w_i \mid \left(w_i > w_j\right) = \widehat{k}_i + \widehat{\boldsymbol{\beta}}_i \mathbf{X}_i + \widehat{\rho}_i \widehat{\sqrt{\sigma_i}} \widehat{\lambda}_i + v_i \tag{26}$$

where $\widehat{\lambda}_i$ is the estimated inverse Mills ratio, v_i is a zero-mean error term, and where

$$\widehat{\boldsymbol{\beta}}_i \mathbf{X}_i = \beta^{educ} educ_i + \beta_1 S_{\exp,i} + \beta_2 S_{\exp,i}^2 \tag{27}$$

and $S_{\exp,i}$ is experience.

I use the estimates of Table 3, with the results for two cross-sections in the years 1981 and 1987 (as results are qualitatively the same across years, I use only these two years). In what follows, I discuss the relations of the distributions included in the wage equation (26) across the two locations, local and Israel.

7.1 Log Hourly Wages

Consider the log hourly wage conditional on selection, $\ln w_i \mid \left(w_i > w_j\right)$. Figure 3 depicts the kernel density (using the Epanechnikov method) and the empirical CDF, for 1981 (panel a) and for 1987 (panel b) in the two locations. Below the graphs is a table which tests for stochastic dominance across locations for 1981 and 1987. The test follows the methodology of Davidson and Duclos (2000) and its implementation by Araar (2006).[25] This is a numerical approach, whereby there is a comparison of the CDF in each location for a range of wages with fixed steps. The idea is to test stochastic dominance (SD) over a restricted range of the distribution. It looks for the longest interval whereby SD obtains. The method consists in estimating $F_i - F_j$ over the range and calculating confidence intervals. The table describes these ranges (w^-, w^+) where the distribution in location i dominates the distribution in location j at first order so that $F_i(\varsigma) > F_j(\varsigma)$ over the range $\varsigma \in [w^-, w^+]$ with 95% confidence. This involves estimation of the points at which there is a reversal of the ranking of the curves, or the intersection points of the dominance curves. The last row in the table indicates the dominance from the last intersection upwards in terms of wages.

Figure 3

The figure suggests that there is no uniform dominance relationship; at lower wages, Israel employment dominates, and, at higher wages, local employment dominates.

I turn now to look at the components of the RHS of equation (26).

7.2 Log Hourly Wages Net of Productivity

If we deduct $\widehat{k}_i$ from log hourly wages we get the following distributions, repeating Figure 3 with this shift.

Figure 4

The local distribution is now well to the right of the Israel employment distribution and dominates it completely from a certain wage upward in 1981 and over the entire range in 1987.

7.3 Skills

I now consider the term $\widehat{\boldsymbol{\beta}}_i \mathbf{X}_i$, elaborated in equation (27), expressing skills and their returns.

7.3.1 Skills Distributions

Figure 5a shows a boxplot of education and experience, the elements of $\mathbf{X}$, across locations.

Figure 5a

[25] Using the dompov module within DASP Version 2.3 in Stata.

The graphs show that workers employed locally have skills with higher mean, median, interquartile range, and variance relative to the Israeli employed ones.

7.3.2 Stochastic Dominance

Next, Figure 5b depicts the kernel density, the empirical CDF, for 1981 (panel a) and for 1987 (panel b) for the expression $\widehat{\boldsymbol{\beta}}_i \mathbf{X}_i = \beta^{educ} educ_i + \beta_1 S_{\exp,i} + \beta_2 S^2_{\exp,i}$, and tests for stochastic dominance across locations for 1981 and 1987, using the same methodology described above.

Figure 5b

Figure 5b suggests that the local distribution dominates the Israel employment one. This is consistent with Borjas, Kauppinen, and Poutvaara (2019) who show that the skill distribution for stayers stochastically dominates the distribution for movers when the rate of return to observable skills is higher locally.[26]Table 3 above has shown that this is the case here. It should be noted that Borjas, Kauppinen, and Poutvaara (2019) also 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 of Table 3 indicate that there is no stochastic dominance in unobservable skills, given the afore-cited low correlation of unobserved skills (ρ_{ij}) across locations.

7.3.3 Wages as a Function of Skills

Using the estimates of Table 3 for 1981 (in panel a) and for 1987 (in panel b), Figure 5c plots the following equation:

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

where

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

and $Ev_i = 0$.

This is a 3D 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 $\widehat{k}_i + \beta_{i,0}$, and with $Ev_i = 0$. Thus, it is a plot of conditional log wages, education, and experience on the axes.

[26]See their page 150 and equation 12.

Figure 5c

The graphs show that at low levels of education and experience, the blue plane, representing log wages of workers who chose to work in Israel, lies above the green plane, representing log wages of workers, who chose local employment. The positions are reversed at high levels of education and experience. The demarcation line, where the switch occurs, is denoted by the dashed line.

Note that workers actually choose on the basis of comparing wages, which are not conditioned by $E\left(\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\widehat{\lambda}_i\right)$ and which do include $v_{i.}$ Thus the green and blue planes do not describe actual wages. Rather, it is a graphical depiction of the following features:

(i) the blue plane is relatively flat, reflecting the relatively low returns to skills in Israel employment

(ii) the blue plane is "pushed up" vertically along the $\ln w$ axis because of the higher $\widehat{k}_i$ term in Israel; for values of skills below the dashed line this makes predicted wages higher in Israel; for values of skills above the dashed line this does not suffice, and predicted wages are higher locally.

To get a sense of the magnitudes embodied in the figure, the table below the graphs shows log wages for each location as predicted by equation (28), as well as the part predicted by average skills in the locality. It does so for 1981 and for 1987. The values of the first row in the table are marked in the graphs of Figure 5c by I (on the blue plane) and L (on the green plane), for Israel and local, respectively. 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 it is the workers employed in Israel who have higher predicted wages (first row) because of the $\widehat{k}_i$ term (contained in C_i).[27] These differences in favor of workers in Israel amount to 0.06 log points in 1981 and to 0.15 log points in 1987.

7.4 Selection

Figure 6 repeats the afore-going analysis for the selection term $\widehat{\rho}_i\widehat{\sqrt{\sigma_i}}\widehat{\lambda}_i$.

Figure 6

Note that $\widehat{\lambda}_i$ is a monotone decreasing function of the probability that an observation is selected into the sample. For the most part, the local distribution dominates the Israel one and is much more dispersed. This

[27] These two points indeed lie below the dashed line.

term, then, contributes to local wages more than to wages in Israel in terms of both first and second moments, but is much smaller in magnitude than skills.

7.5 Tasks and Unobserved Skills

A different take on the distributions is provided by looking at the relation between tasks in the two locations. The analysis above yields:[28]

$$\begin{aligned}
\ln t_{Israel} &= \mu_{Israel} + \frac{\sigma_{local,Israel}}{\sigma_{local}}(\ln t_{local} - \mu_{local}) + \varepsilon_{Israel} \qquad (29)\\
&= \left(\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}\right) + \frac{\sigma_{local,Israel}}{\sigma_{local}} \ln t_{local} + \varepsilon_{Israel}
\end{aligned}$$

where:

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

Figure 7 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 7

The figure has 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 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 $\mu_{Israel} - \frac{\sigma_{local,Israel}}{\sigma_{local}}\mu_{local}$,[29] 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 in orange. The data points are distributed – conditional on the $\ln t_{local}$ value – with $var\ \varepsilon_{Israel}$. The black line in the figure is the 45 degree line serving as the line of equal income ($\ln w_{local} = \ln w_{Israel}$).[30]This 45

[28]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 $\ln t_{Israel}$.

[29]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 $\beta_0 = 0$.

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

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. The green and blue lines express the average $\ln t_i$ values for local and Israel employment, respectively.

Three major features of the analysis are manifested in the figure.

The effect of the move to the rich economy. 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.[31]

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 5c.

Positive selection on unobservables. The figure illustrates the positive selection on unobservables in each location.[32]In 1981, the term $\frac{\sigma_{local,Israel}}{\sigma_{local}}$ is positive and less than 1, the case of comparative advantage discussed above in sub-section 6.2. 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, the case of absolute advantage discussed above. 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.

Comparative statics and policy effects. One question of interest is to consider how moving behavior would change following changes in the observed skill premia and in the unobserved skills distributions. The model is able to predict the size of moving when key parameters (π, μ), determining first moments, change. But changes in second moments $(\sigma_{ii}, \sigma_{ij})$ lead to ambiguous outcomes, as contradictory effects are at play. These results can be seen in the graphical framework of Figure 7 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

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

[32] This means that in each sector

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

choose the local economy and more move. This is the effect of the productivity element discussed above (and again below).

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. This is an expression of the task rewards element.

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.

The last three changes could be generated by changes in task demanded across locations. This analysis also 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.

7.6 Summing Up

The afore-going analysis depicts a similar qualitative picture to the one gleaned from the means analysis of the preceding section. The emerging picture is that the productivity differential shifts the Israel wage distribution significantly to the right; were it not for this differential, the local wage distribution dominates the Israel one due to both skills (for the main part) and selection, and is much more dispersed. When the productivity differential is included, the wage distributions overlap, albeit with greater dispersion and with higher kurtosis (fat tails) for the local distribution.

8 Broader Contexts

I turn to discuss two issues that place the findings here in broader contexts.

8.1 Implications for Development Accounting

The essential point of linkage between the issues debated in the development accounting literature and the current paper is that both make the distinction between what characterizes rich and poor economies in terms of technology, capital, and institutions, and what constitutes human capital, embodied in people. I therefore divide the following discussion into these two aspects.

8.1.1 The Role of Technology, Capital, and Institutions

I have defined the variable z as follows:

$$\begin{aligned} z_i &\equiv K_i^{\alpha} A_i^{1-\alpha} T_i^{-\alpha} \\ &= \left(\frac{K_i}{T_i}\right)^{\alpha} A_i^{1-\alpha} \end{aligned}$$

This variable captures the role of technology, capital, and institutions. Using equation (15), z differences across locations are given by:

$$\ln z_i - \ln z_j = \ln \pi_i - \ln \pi_j \tag{30}$$

I have used the estimates of the wage equation reported in Table 5 which 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 z differentials. The estimates of $\widehat{k}_i - \widehat{k}_j$ were shown to 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.

In the development accounting literature, the analysis of Hendricks and Schoellman (2018) breaks down the differential of GDP per capita across countries into a wage differential capturing a country differential and a human capital differential. Their analysis (see their pages 670-672) postulates the following accounting relations:

$$\ln y_c - \ln y_{c'} = \ln z_c - \ln z_{c'} + \ln h_c - \ln h_{c'} \tag{31}$$

where c, c' denote two different countries, y is GDP per capita, and h is human capital per worker. Their z_c is defined as

$$z_c \equiv \left(\frac{K_c}{Y_c}\right)^{\frac{\alpha}{1-\alpha}} A_c \tag{32}$$

where K_c, Y_c, A_c are the capital, output, and technology of country c, respectively. It captures similar elements to the z variable in the current model, with the important distinction that here z pertains to a location-task product and in their case it refers to the GDP of the entire economy.

They call the ratio $\frac{\ln z_c - \ln z_{c'}}{\ln y_c - \ln y_{c'}}$ the country share and the ratio $\frac{\ln h_c - \ln h_{c'}}{\ln y_c - \ln y_{c'}}$ the human capital share in the GDP per capita differential.

Postulating worker i wages as

$$\ln w_{i,c} = \ln(1-\alpha)z_c + \ln h_i \tag{33}$$

they get that the country share is therefore given by:

$$\frac{\ln w_{i,c} - \ln w_{i,c'}}{\ln y_c - \ln y_{c'}} = \frac{\ln z_c - \ln z_{c'}}{\ln y_c - \ln y_{c'}} \tag{34}$$

Hendricks and Schoellman (2018) then use data on these variables across countries, comparing each country to the U.S. Using wage differentials of immigrants pre- and post-migration to compute $\ln w_{i,c} - \ln w_{i,U.S.}$ they report (see their Table II) country shares ranging between 0.34 and 0.52; summing over different empirical checks they point to 0.40 as the country share. The values of the $\frac{z_{U.S.}}{z_c}$ ratios (same table) range between 1.8 and 3.2. This is a very similar range to the one estimated in the current paper for the z ratio across locations, as reported above, namely 1.6 to 3.

8.1.2 The Role of Human Capital

For Hendricks and Schoellman (2018), the human capital share in the GDP per capita differential is simply the complement of the country share discussed above. Hence their results range between 0.48 and 0.66. For the development accounting literature this result is important, as it assigns a substantial role to human capital differences, higher than the one typically assumed previously.

The current paper does not estimate human capital differences across countries, as it looks at wage differentials of workers who are stayers and movers from one single economy, the Palestinian one. What this paper does show is that in terms of human capital tasks, there is a big offset effect. The total wage differential across locations ranges between -0.08 and $+0.17$ log points only. This is so despite the big z differential in favor of the Israeli economy. The offset comes through the task term, the $\widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel}$ difference, which ranges between 0.60 and 1 log points.

8.1.3 The Similarities and Differences Across Studies

The preceding discussion makes it clear that there are important differences between the methodology and the results of the afore-cited HS (2018, 2019) studies, which are key in the DA literature, and those of this paper. In what follows I discuss the differences and point out the similarities.

Data used. 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 of Palestinian men in the period 1981-1987 and pertain to current year jobs.

Inference of the z differentials. The methodology used is quite different. HS (2018, 2019) use GDP per worker across countries and deduce estimates of z differences in an accounting exercise: they compare 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, not to GDP.

Wage differentials. 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 afore-cited wages, across the U.S. and source countries. Equations (31) – (34) above summarize their analysis. The findings of HS (2018) were reported above; they remains broadly true (the cited range turns into 50% to 75%) when, in HS (2019), they take into account a whole host of factors, in particular, imperfect substitution of skills in production and endogenous choice of production.[33]

The current paper has tasks, rather than human capital stocks per se, producing a good in a particular location, and the latter output is not aggregate output, 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 economy. The task bundle reflects the worker (his skills, $\mathbf{X}$ and his job task returns ($\boldsymbol{\beta}$)).

It is clear, then, that the wage differential examined here, across locations, is not the same as the HS wage differential. The wage differential

[33]See Tables IV,V and VIII in HS (2018) and Tables 7, 10, 11, 13, and 14 in HS (2019).

here reflects both the z cross-country differential, *as* in HS, as well as the task differential across locations, which reflects worker skills and job task returns, *unlike* the approach of HS.

Similarities. It can be conjectured that the HS results may still hold true in the current case. This is based on two points: first, as noted, the z differences are similar, the differences in definitions and computation methodologies notwithstanding; second, 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, or more, here. The second point, however, is not examined in the current paper. Likewise, the findings here, whereby the foreign task bundle has a relatively 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. However, 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.

Selection effects. HS (2018) also examine the effects of selection. In their data, they find positive selection on observables and on unobservables (see their Figure II) while the current paper finds negative selection on observables, which plays an important role quantitatively, and positive selection on unobservables. It is notable, though, that HS 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 sectorial composition of employment. This last point is inherent and fundamental in the analysis undertaken here.

8.2 Applicability to Other Cases of Movers

The analysis here is potentially pertinent to many cases worldwide. The following is a brief survey of recent papers which indicate that the phenomenon of workers from poor countries working in *manual tasks* in rich countries is very prevalent.

Cassidy (2019) uses data on men aged 25-64 from the US census Integrated Public Use Microdata Series in the period 1970-2010 and the US Department of Labor's O*NET database. His key findings are that immigrants have on average higher manual and lower analytical and interactive task requirements than natives, and this gap has expanded greatly over the past several decades. His Tables 3 and 4 puts the gap between native and immigrant in terms of manual task requirements at between 0.05 and 0.20, with most estimates between 0.08 and 0.11, depending upon the year examined, and conditioning on education, age, or time since migration. An earlier study with similar data covering the period 1960 to 2000, by Peri and Sparber (2009), found that foreign-born workers specialize in occupations intensive in manual and physical labor skills while natives hold jobs more intensive in communication-language tasks. Lewis and Peri (2015) report further results in this direction and highlight the mechanism whereby migrant workers fill manual-intensive jobs that are often at the bottom of the career ladder for natives; hence in locations with large inflows of immigrants, native workers move more rapidly toward communication-intensive and more complex type of jobs. Consistently with these findings, Card and Raphael (2013) document relatively high poverty rates amongst immigrants to the U.S., particularly those from poor economies.

Dustmann and Frattini (2013) document sizable differences in educational attainment between the foreign and native born in most Western European nations, with immigrants considerably less educated than the native born. They are occupationally segregated from the native born, working in lower paying, less prestigious occupational categories. They are also considerably less likely to be employed and considerably more likely to have earnings in the lower deciles of the earnings distribution of the host country.

The afore-cited survey by Peri (2016) stresses the importance of these patterns. He suggests that manual abilities are transferable across countries but other abilities, such as communication abilities (especially if languages differ), are much harder to transfer.

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. This claim ties in with key discussions 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 may play contradictory roles, yielding lower net gains.

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. Movers and stayers are not performing the same tasks in the home and host countries. Importantly, they face bundles of location-tasks-wages. This has important repercussions in terms of the rewards to skills which they get and, as the analysis here demonstrates, in terms of the movers' wage gains. The analysis of Autor and Handel (2013) which explicitly examines wages, jobs, and tasks within the framework of a similar self-selection model, is of particular importance. It indicates that this task selection model is an empirically-relevant one and points to negative relationships between returns to analytical and manual tasks.

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.

The challenge for future research is to get the necessary data so as undertake similar decompositions in prevalent cases, whereby confounding factors are present, and try to disentangle their relative, and potentially contradictory, effects. It has been shown here that the model to be studied in these cases should cater for multiple determinants in order to avoid mis-specification, and would thus need a very rich data set.

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 highly pertinent in the current context. Foreign and home tasks requirements undergo changes, and so task requirements of movers and stayers change. Hence a task-based approach is crucial in terms of studying 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] Angrist, Joshua D. , 1996. "Short-Run Demand for Palestinian Labor," **Journal of Labor Economics** 14, 3, 425-453.

[6] Angrist, Joshua D., and Jörn-Steffen Pischke, 2009. **Mostly Harmless Econometrics: An Empiricist's Companion**. Princeton University Press, Princeton.

[7] Araar, Abdelkrim, 2006. "Poverty, Inequality and Stochastic Dominance, Theory and Practice: Illustration with Burkina Faso Surveys," WP06-34, CIRPEE.

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

[9] 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.

[10] 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.

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

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

[13] 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.

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

[15] Card, David and Steven Raphael, 2013. "Introduction," in David Card and Steven Raphael (eds.) **Immigration, Poverty, and Socioeconomic Inequality**, Chapter 1, 1-26, New York, NY: Russel Sage Foundation.

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

[17] Caselli, Francesco, 2016. **Technology Differences over Time and Space.** CREI Lectures, Princeton University Press.

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

[19] Cassidy, Hugh, 2019. "Occupational Attainment of Natives and Immigrants: A Cross-Cohort Analysis," **Journal of Human Capital** 13, 3, 375-409.

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

[21] 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.

[22] Davidson, Russell and Jean-Yves Duclos, 2000. "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," **Econometrica** 68, 6, 1435-1464.

[23] D'Haultfoeuille, Xavier and Arnaud Maurel, 2013. "Inference on an Extended Roy Model, with an Application to Schooling Decisions in France," **Journal of Econometrics 1**74, 95–106

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

[25] D'Haultfoeuille, Xavier, Arnaud Maurel, Xiaoyun Qiu, and Yichong Zhang, 2019. "Estimating Selection Models without Instrument with Stata," NBER Working Paper No. 25823.

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

[27] Gathmann, Christina and Uta Schönberg, 2010. "How General Is Human Capital? A Task-Based Approach," **Journal of Labor Economics,** 28,1, 1-48.

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

[29] 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.

[30] 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.

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

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

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

[34] 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.

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

[36] 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.

[37] 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.

[38] Lewis, Ethan and Giovanni Peri, 2015. "Immigration and the Economyof Cities and Regions," Chapter 10 in G.Duranton, J. Vernon Henderson, W.C. Strange (eds.) **Handbook of Regional and Urban Economic**s Volume5A, 625-685, Elsevier, Amsterdam.

[39] OECD, 2020. Permanent immigrant inflows (indicator). doi: 10.1787/304546b6-en.

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

[41] Peri, Giovanni, and Chad Sparber. 2009. "Task Specialization, Immigration, and Wages," **American Economic Journal: Applied Economics** 1(3): 135–69.

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

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

[44] UN, 2019. Workbook: UNMigrant Stock by Origin and Destimation 2019, Excel file, downloaded from

https://www.un.org/en/development/desa/population/migration/data/estimates2/estima

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

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

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

10 Appendix: Econometric Methodologies

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

10.1 The 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:

$$\ln w_i = \ln \pi_i + \beta_i \mathbf{X} + u_i, \tag{35}$$

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

When estimating equation (35), I take into account sample selection, which is inherent in the model. Thus define the variable z^* :

$$\begin{aligned} z^* &= \ln w_i + \ln(1 - k_i(\mathbf{L})) + \ln(1 - \gamma_i) - \ln w_j - \ln(1 - k_j(\mathbf{L})) - \ln(1 - \gamma_j) \\ &= \ln \pi_i - \ln \pi_j \\ &\quad + \ln(1 - k_i(\mathbf{L})) - \ln(1 - k_j(\mathbf{L})) \\ &\quad + \ln(1 - \gamma_i) - \ln(1 - \gamma_j) \\ &\quad + \beta_i \mathbf{X} - \beta_j \mathbf{X} \\ &\quad + u_i - u_j \end{aligned} \tag{36}$$

and the indicator variable z :

$$\begin{aligned} z &= 1 \; if \; z^* > 0 \\ z &= 0 \text{ otherwise} \end{aligned} \tag{37}$$

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

$$\begin{aligned} \Pr(z &= 1) = \Phi(\ln \frac{\pi_i}{\pi_j} + \ln \frac{(1 - k_i(\mathbf{L}))}{(1 - k_j(\mathbf{L}))} + \ln(1 - \gamma_i) - \ln(1 - \gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \\ \Pr(z &= 0) = 1 - \Phi(\ln \frac{\pi_i}{\pi_j} + \ln \frac{(1 - k_i(\mathbf{L}))}{(1 - k_j(\mathbf{L}))} + \ln(1 - \gamma_i) - \ln(1 - \gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \end{aligned} \tag{38}$$

The observed $\ln w_i$ is given by:

$$\ln w_i \mid (z = 1) = \ln \pi_i + \beta_i \mathbf{X} + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*} \right] \lambda(c_i) + v_i \tag{39}$$

where v_i is a zero mean error uncorrelated with the regressors and 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 $\phi(\cdot), \Phi(\cdot)$ denoting the density and CDF of a standard normal variable, respectively.

This may also be written as follows:

$$\ln w_i \mid (z=1) = \ln \pi_i + \beta_i \mathbf{X} + \rho_i \sqrt{\sigma_{i\,i}}\, \lambda\,(c_i\,) + v_i \tag{40}$$

A similar equation holds true for the other location. Note that while the **X** vector appears in both (38) and (40), the **L** vector appears only in the selection equation (38). I estimate the model using Heckman's (1979) two-step consistent estimation procedure. One can interpret the selection bias in (**??**) 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**:

$$\frac{\partial E(\ln w_i \mid (z=1))}{\partial x_k} = \beta_{ik} + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}\right] \frac{\partial \lambda}{\partial c_i} \frac{\partial c_i}{\partial x_k} \tag{41}$$

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 5.1.1.

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 (38). 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[34]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:

$$\begin{aligned}\Pr(z &= 1) = \Phi(\ln \frac{\pi_i}{\pi_j} + \sum_p \theta_p \cdot l_p^j - \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^j - \sum_n \gamma_n Y_n^i \qquad (42) \\ &\quad + \ln(1-\gamma_i) - \ln(1-\gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j) \\ \Pr(z &= 0) = 1 - \Phi(\ln \frac{\pi_i}{\pi_j} + \sum_p \theta_p \cdot l_p^j - \sum_p \theta_p \cdot l_p^i + \sum_n \gamma_n Y_n^j - \sum_n \gamma_n Y_n^i \\ &\quad + \ln(1-\gamma_i) - \ln(1-\gamma_j) + \beta_i \mathbf{X} - \beta_j \mathbf{X} + u_i - u_j)\end{aligned}$$

The estimated wage equation is the following:

$$\begin{aligned}\ln w_i \quad | \quad & \text{location } i = \ln \pi_i + \beta_{0i} + \beta_{1\,i} educ + \beta_{2\,i} exp + \beta_{3i} exp^2 \\ & + \sum_{m=2}^{4} \gamma_m Q_m + \left[\frac{\sigma_{ii} - \sigma_{ij}}{\sigma^*}\right] \lambda\,(c_i) + u_i \qquad (43)\end{aligned}$$

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).[35] Education (*educ*) and experience (*exp*) are defined in years.

The benchmark specification reported in the text [column (1) of Table 2] features the geographical exclusion restrictions. The alternative, specification 2 includes the variables discussed above contained in $\mathbf{L}$, so there are

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

[35]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.

three exclusion restrictions. Specification (3) uses OLS to test for the effect of selection correction (running only the wage equation).

10.2 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.[36]

The rationale of their methodology is as follows:[37]

> ... 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:[38]

> 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

[36] 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.

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

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

> 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.

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

11 Tables and Figures

Table 1
Sample Statistics, LFS data
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.

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

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

Table 2: Heckman Two Step Estimates 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)
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. According to the model one observes $\ln w_i$ only if $z^* > 0$ i.e., when $z = 1$. So we have:

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

Thus:

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

2. The wage equation in panel b is given by:

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

It is estimated with two sets of exclusion restrictions in columns 1 and 2, respectively, and uses OLS in column 3 (dropping $\lambda\,(c_i\,))$.

3. For the exclusion restrictions, Set 1 is given by

$$
\mathbf{L} \in [\text{region of residence}]
$$

Set 2 is given by

$$
\mathbf{L} \in \; [\text{region of residence, marital status, urban status}\;]
$$

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

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

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

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

8. The second moments satisfy the following relation:

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

Table 3: Heckman Two Step Estimates
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)
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) Local1984	(8) 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)
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)
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)
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, 2, and 4-7 in Table 2.
2. Panel b uses set 1 for the exclusion restrictions given by

$$\mathbf{L} \in [\text{region of residence}]$$

The first stage is reported in panel a.

3. The second moments satisfy the following relations:

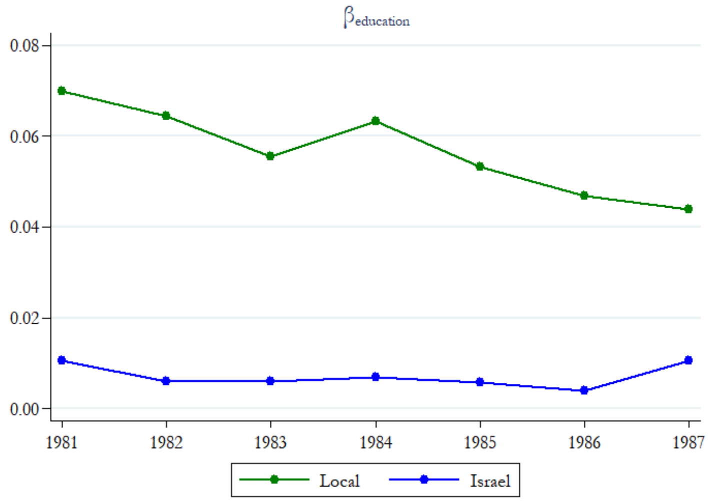
$$
\begin{aligned}
\rho_i &= \frac{\sigma_{ii} - \sigma_{ij}}{\sqrt{\sigma_{ii}}\sigma^*} \\
\rho_j &= \frac{\sigma_{jj} - \sigma_{ij}}{\sqrt{\sigma_{jj}}\sigma^*} \\
\rho_{ij} &= \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}
\end{aligned}
$$

Hence:

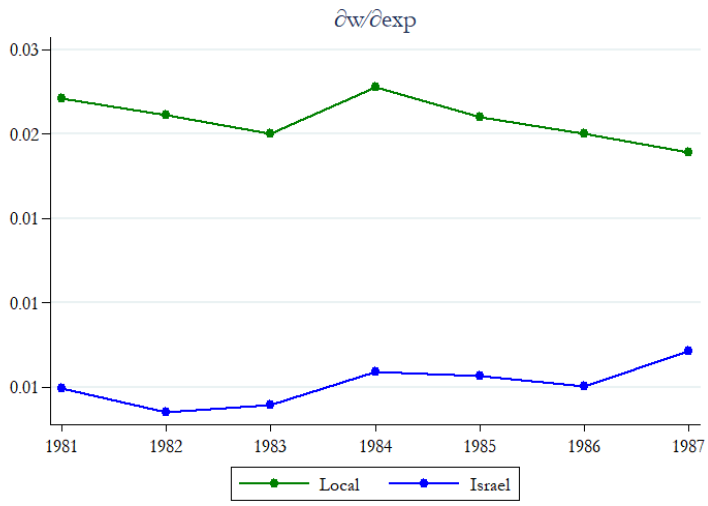
$$
\begin{aligned}
\frac{\rho_i}{\rho_j} &= \frac{\sigma_{ii} - \sigma_{ij}}{\sqrt{\sigma_{ii}}\sigma^*} \frac{\sqrt{\sigma_{jj}}\sigma^*}{\sigma_{jj} - \sigma_{ij}} \\
&= \frac{\sqrt{\sigma_{jj}}}{\sqrt{\sigma_{ii}}} \frac{\sigma_{ii} - \sigma_{ij}}{\sigma_{jj} - \sigma_{ij}}
\end{aligned}
$$

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.

Figure 1: Point Estimates of Skills Returns



a. Returns to education



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

Notes:
Based on the estimates reported in Table 3.

Table 4 : Heckman and Semi Parametric Estimates 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. The Heckman estimates are taken from Table 3.

2.The semi-parametric estimation methodology is described in sub-section 5.1.2 and in the Appendix.

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

$$\overline{\ln w_{local}} \quad | \quad (w_{local} > w_{Israel}) = \widehat{k}_{local} + \widehat{\boldsymbol{\beta}}_{local}\overline{\mathbf{X}}_{local} + \left(\widehat{\rho}_{local}\widehat{\sqrt{\sigma_{local}}}\right)\overline{\widehat{\lambda_{local}}}$$

$$\overline{\ln w_{Israel}} \quad | \quad (w_{Israel} > w_{local}) = \widehat{k}_{Israel} + \widehat{\boldsymbol{\beta}}_{Israel}\overline{\mathbf{X}}_{Israel} + \left(\widehat{\rho_{Israel}}\widehat{\sqrt{\sigma_{Israel}}}\right)\overline{\widehat{\lambda_{Israel}}}$$

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

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
$\widehat{\rho}\widehat{\sqrt{\sigma}}\overline{\widehat{\lambda}}$	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
$\widehat{\rho}\widehat{\sqrt{\sigma}}\overline{\widehat{\lambda}}$	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
$\widehat{\rho}\widehat{\sqrt{\sigma}}\overline{\widehat{\lambda}}$	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
$\widehat{\rho}\widehat{\sqrt{\sigma}}\overline{\widehat{\lambda}}$	0.01	0.10	-0.09

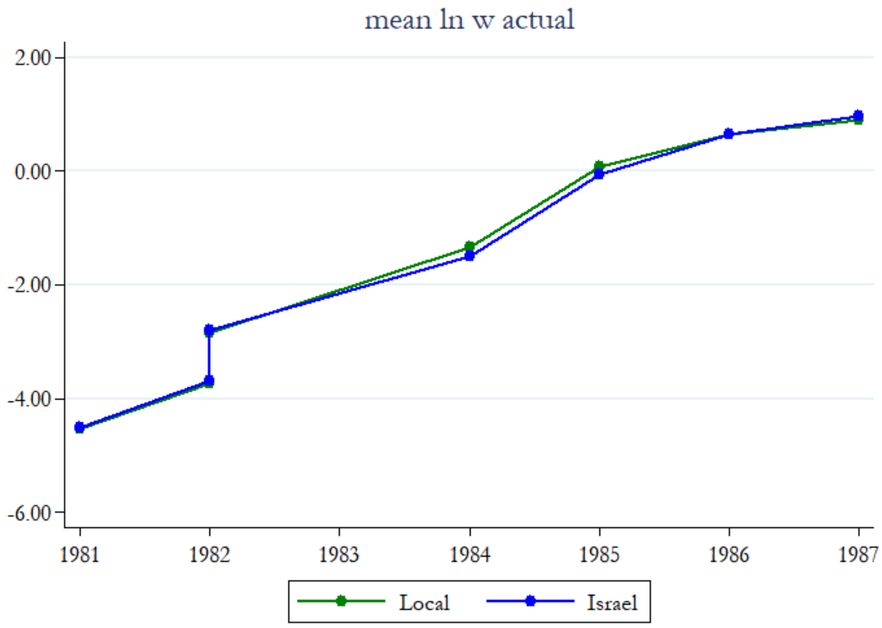
1985	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	0.08	-0.06	0.14
$\widehat{k}$	-0.88	-0.40	-0.48
$\widehat{\beta}\overline{\mathbf{X}}$	1.06	0.26	0.80
$\overline{\mathbf{X}}_{Israel}(\widehat{\beta}_{local} - \widehat{\beta}_{Israel})$			0.71
$\widehat{\beta}_{local}(\overline{\mathbf{X}}_{local} - \overline{\mathbf{X}}_{Israel})$			0.09
$\widehat{\rho}\widehat{\sqrt{\sigma}}\overline{\widehat{\lambda}}$	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
$\widehat{\rho}\widehat{\sqrt{\sigma}}\overline{\widehat{\lambda}}$	0.00	0.03	-0.03

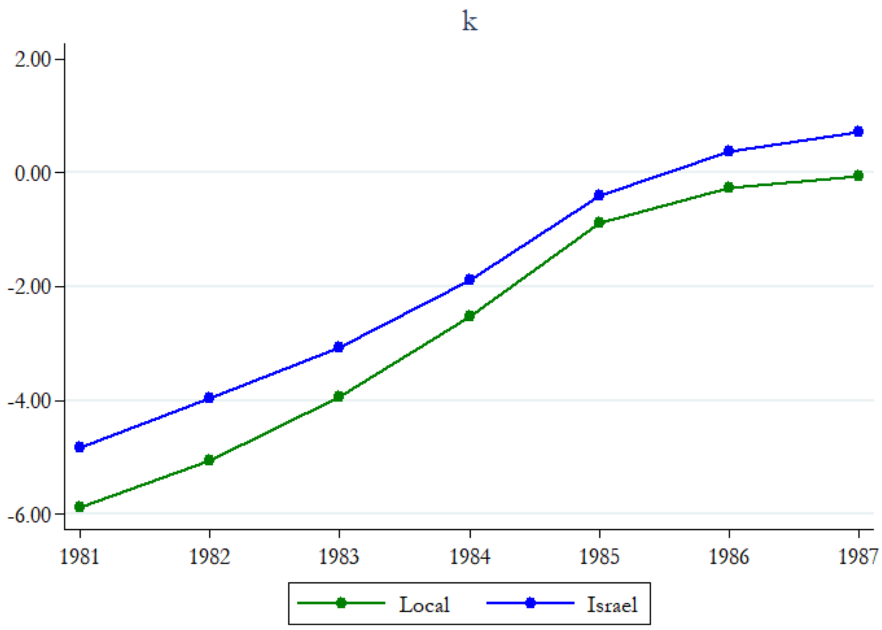
1987	**local**	**Israel**	**difference**
mean $\overline{\ln w}$ actual	0.90	0.97	-0.08
$\widehat{k}$	-0.06	0.72	-0.78
$\widehat{\boldsymbol{\beta}}\overline{\mathbf{X}}$	0.90	0.29	0.61
$\overline{\mathbf{X}}_{Israel}(\widehat{\boldsymbol{\beta}}_{local} - \widehat{\boldsymbol{\beta}}_{Israel})$			0.54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

Notes:
The table is based on the point estimates reported in Table 3.

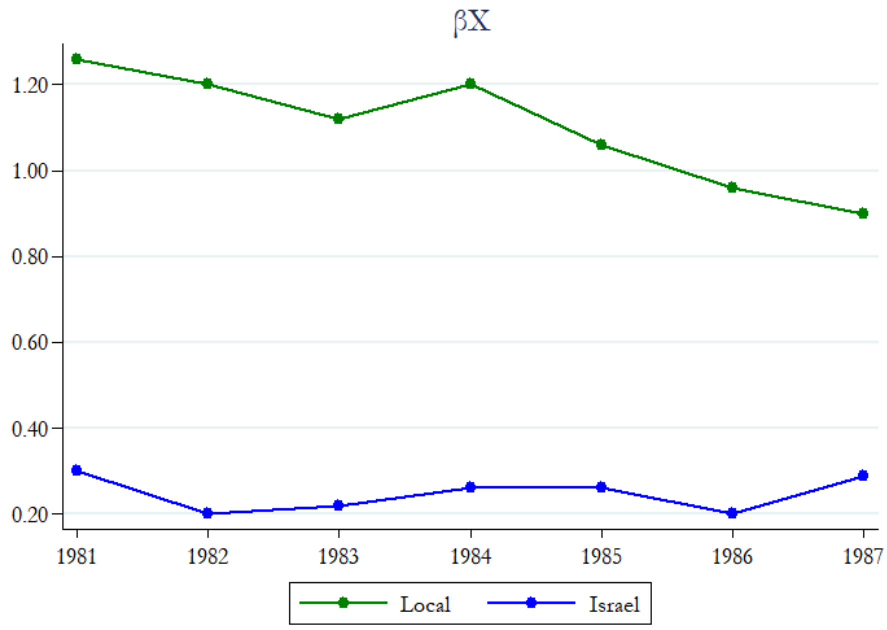
Figure 2: Log Wage Regressions Decompositions



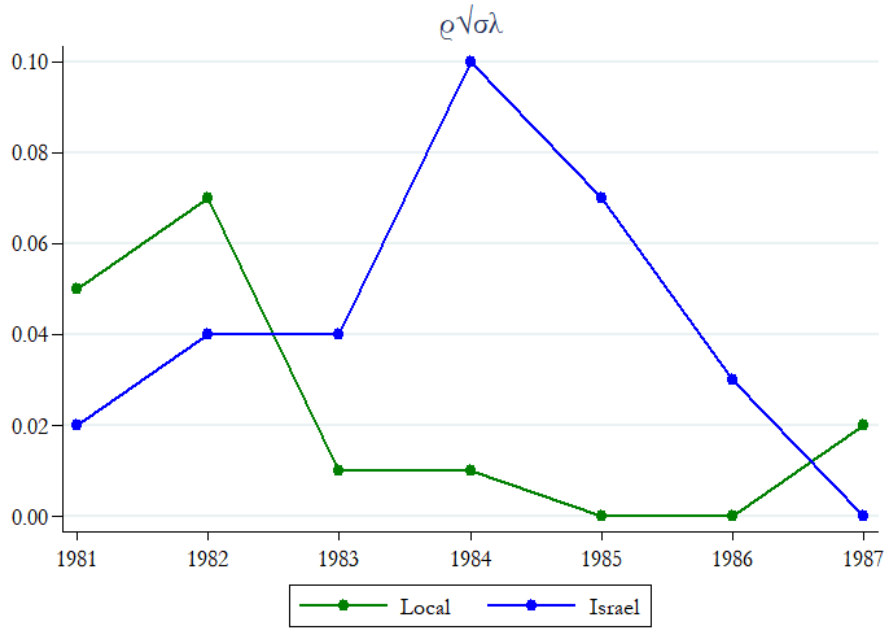
a. **Mean Log Wages**



b. Wage Equation $\widehat{k}$



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



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

Notes:
Based on Table 5.

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%

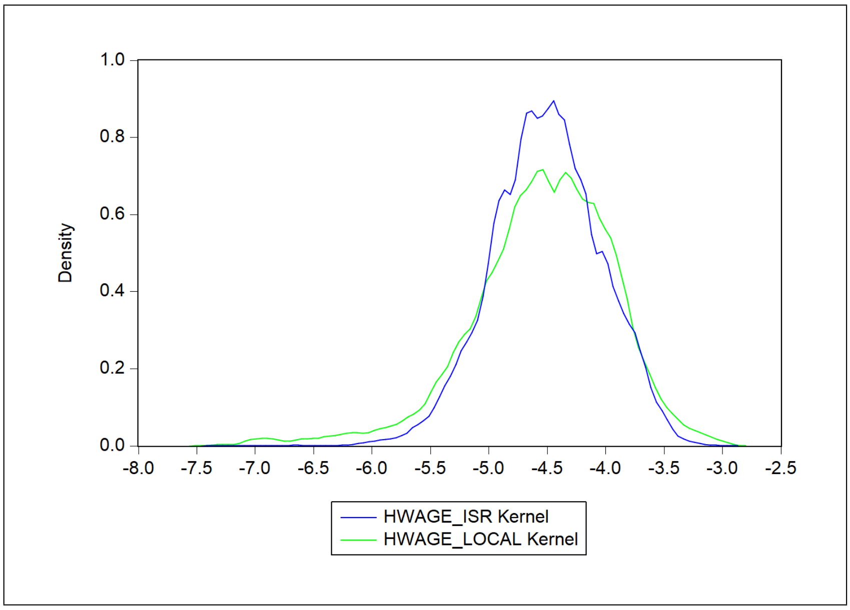
b. Occupation Distributions

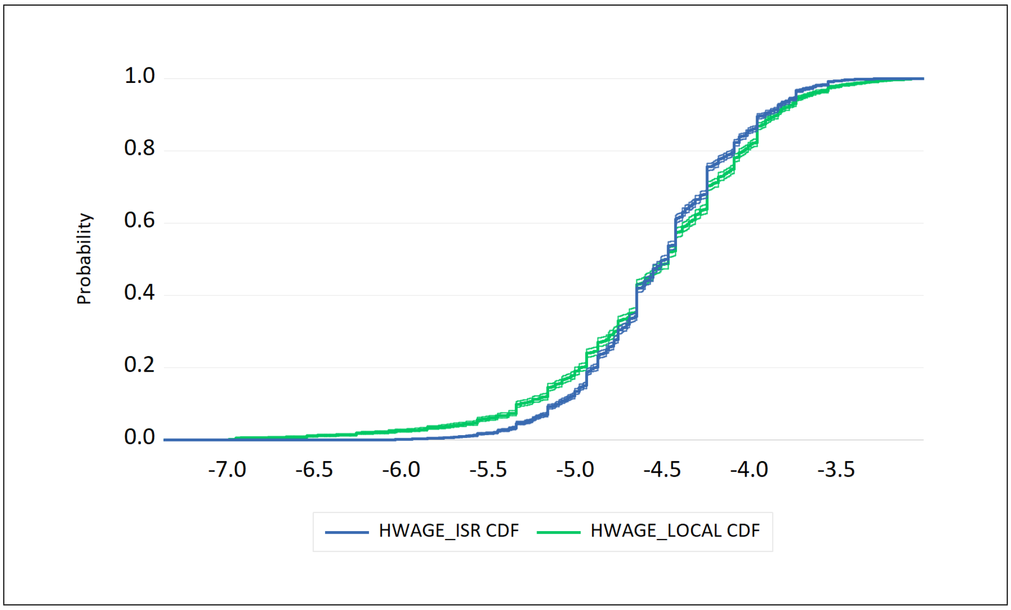
occupation	Local	Israel
academic	6%	0%
professionals	12%	1%
managers	1%	0%
clerical workers	9%	1%
agents, sales and service	12%	14%
skilled job in agriculture	4%	13%
manufacturing and construction skilled jobs	35%	29%
unskilled	22%	42%

Note:
The table refers to data from 1987.

Figure 3: Log Hourly wages

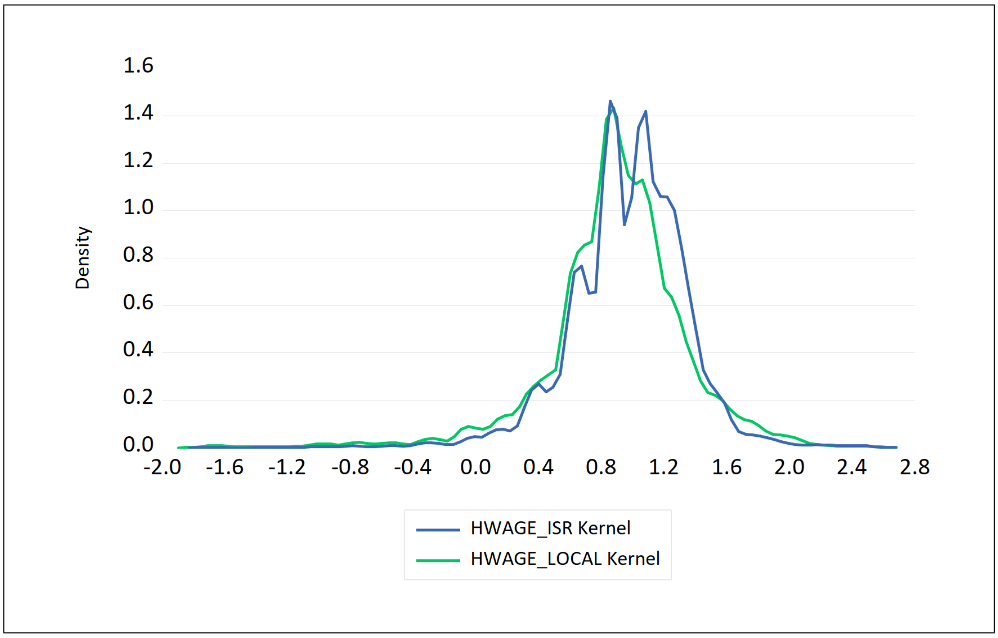
a. 1981

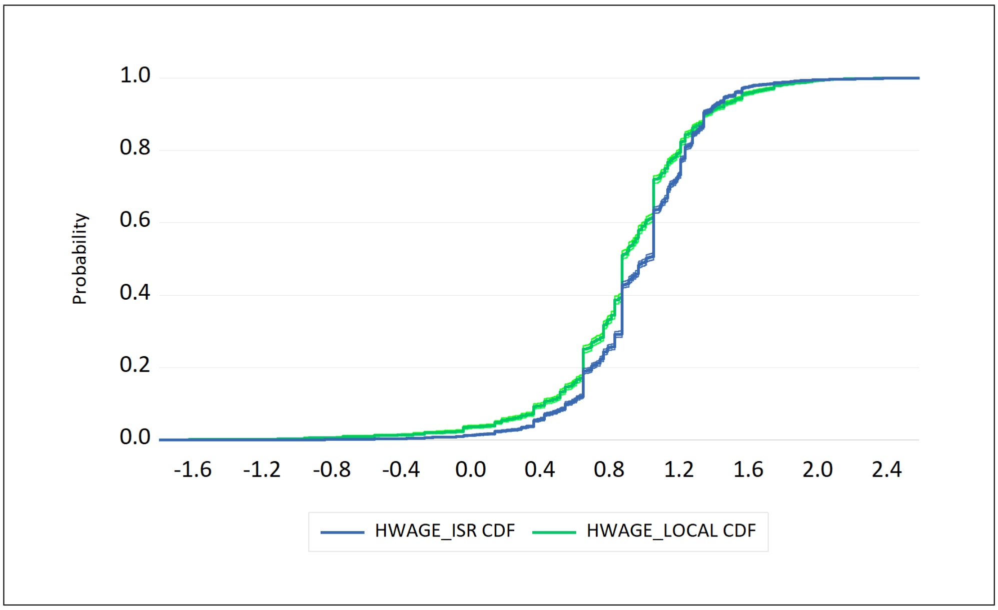




Intersection Number	Intersection Point			Case
	Exact	Min. range	Max. range	
1	.	-4.968	-4.967	B
2	.	-4.966	-4.965	B
3	-4.962	.	.	A
4	.	-4.959	-4.957	B
5	-4.956	.	.	B
6	∞	.	.	A

b. 1987

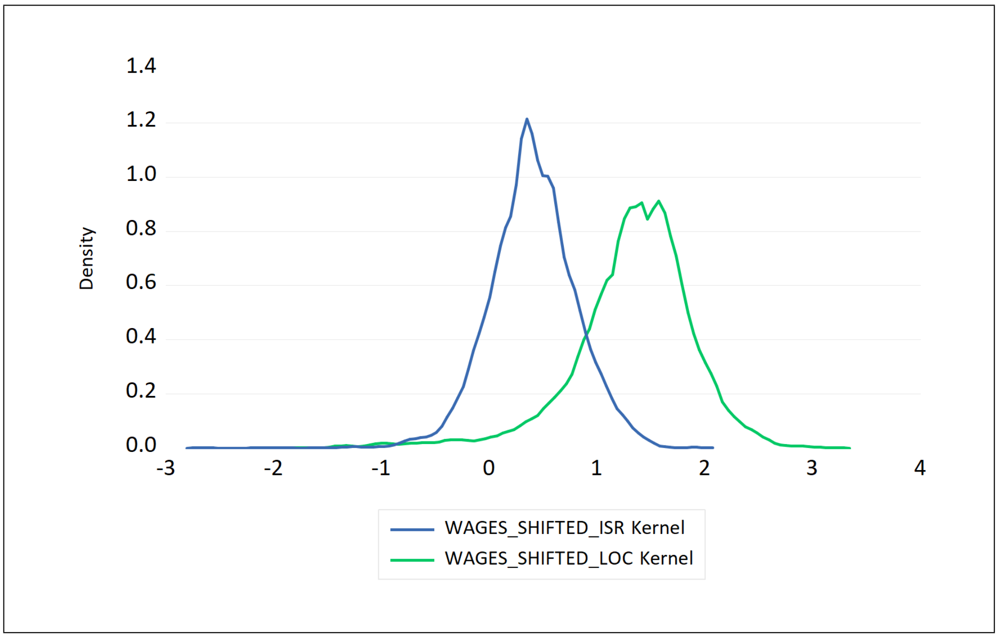


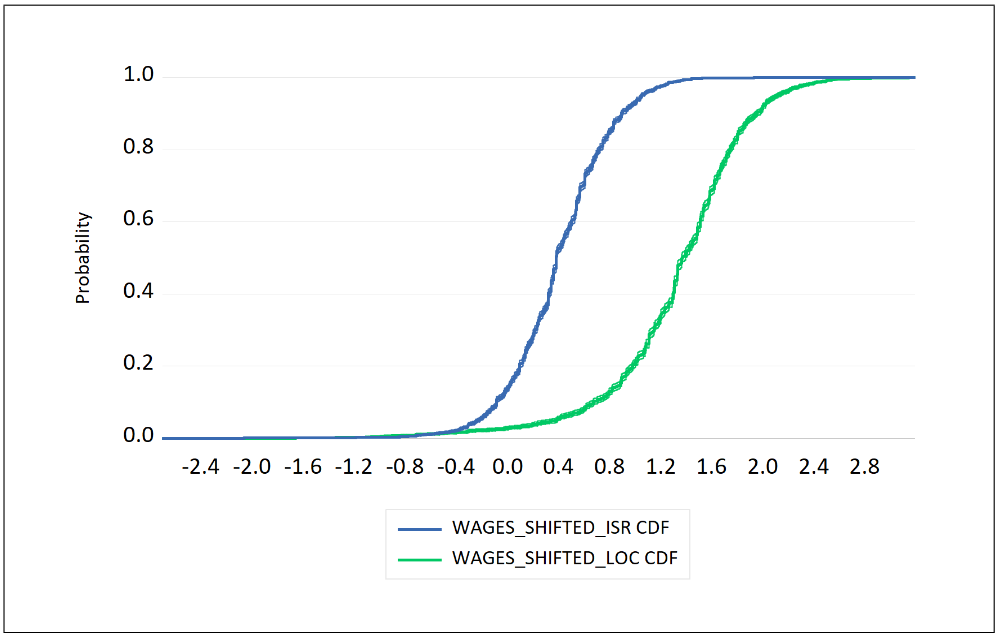


Intersection Number	Intersection Point			Case
	Exact	Min. range	Max. range	
1	0.427	.	.	B
2	∞	.	.	A

Notes :
Case A: Before this intersection, Local dominates Israel
Case B: Before this intersection, Israel dominates Local.

Figure 4: Wages Shifted
a. 1981

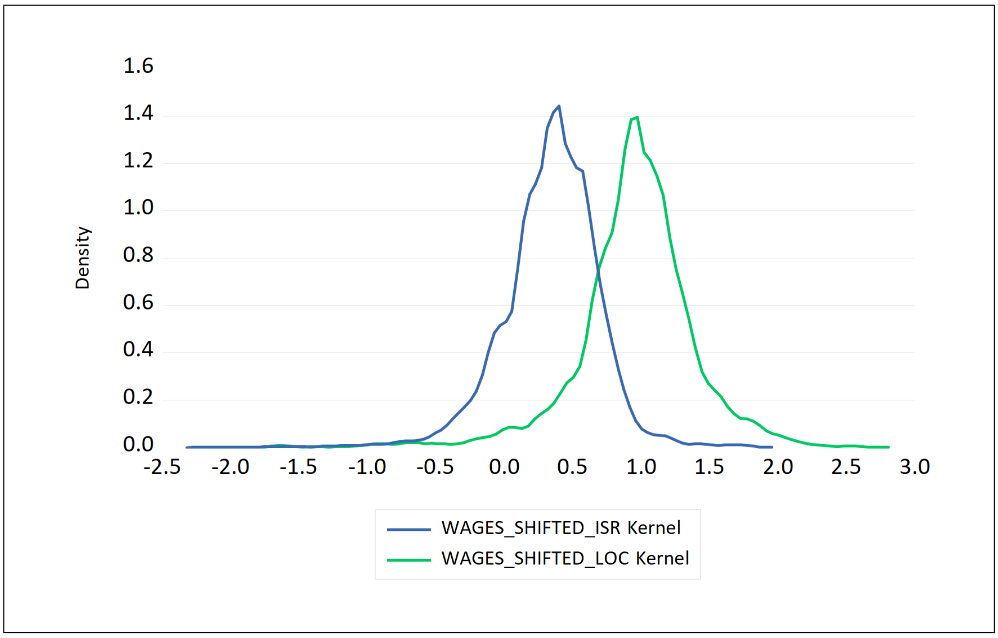


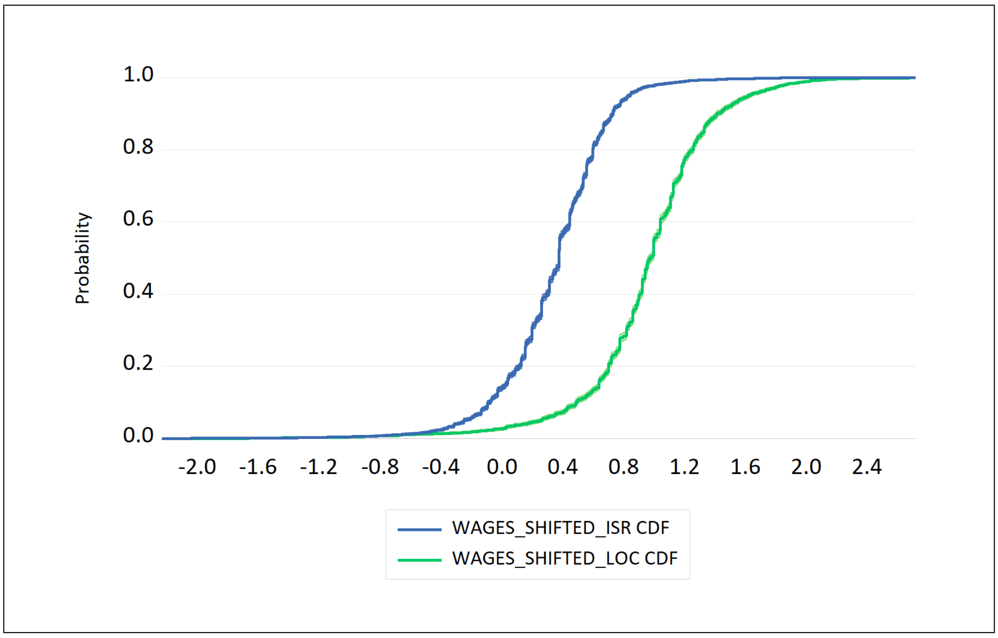


Intersection Number	Intersection Point			Case
	Exact	Min. range	Max. range	
1	-1.375	.	.	A
2	.	-0.751	-0.725	B
3	-0.718	.	.	B
4	∞	.	.	A

Notes :
Case A: Before this intersection, Local dominates Israel
Case B: Before this intersection, Israel dominates Local.

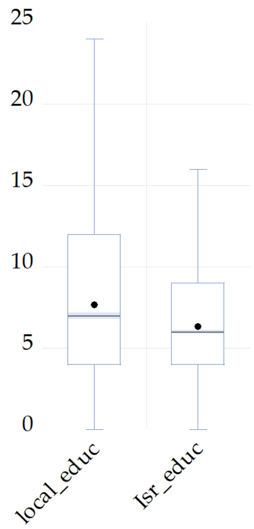
b. 1987



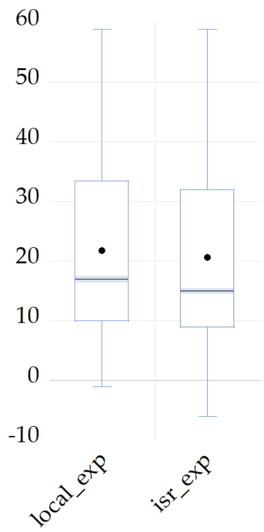


No intersection found. Local dominates Israel.

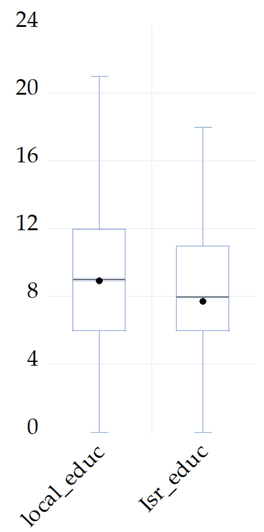
Figure 5a: Education and Experience Boxplots
a. 1981

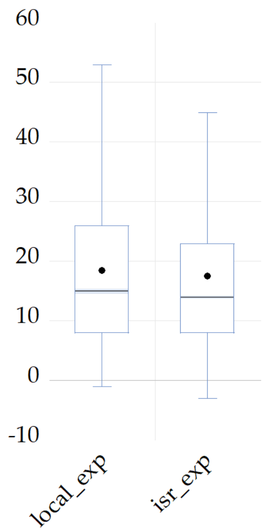


education



b. 1987





Notes:

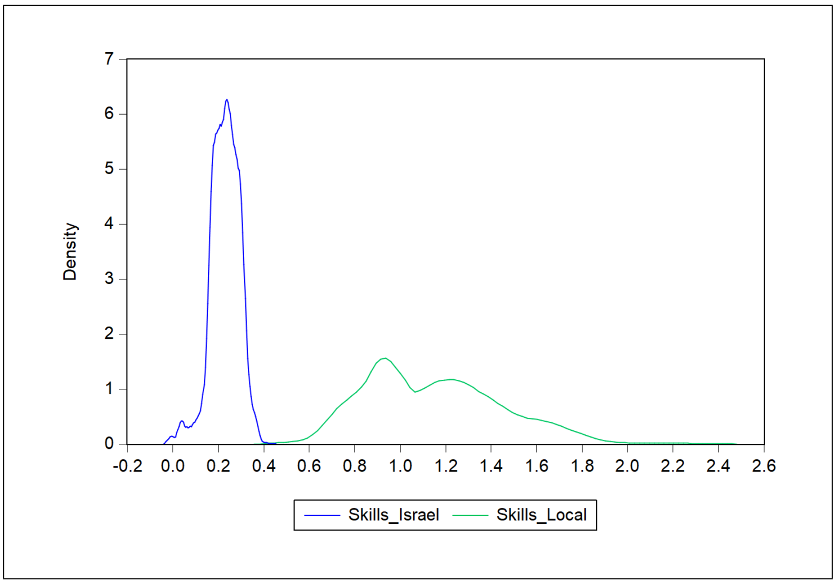
1. The box portion of the boxplot represents the first and third quartiles.
2. The median is depicted in the line through the center of the box, while the mean is drawn as a point.

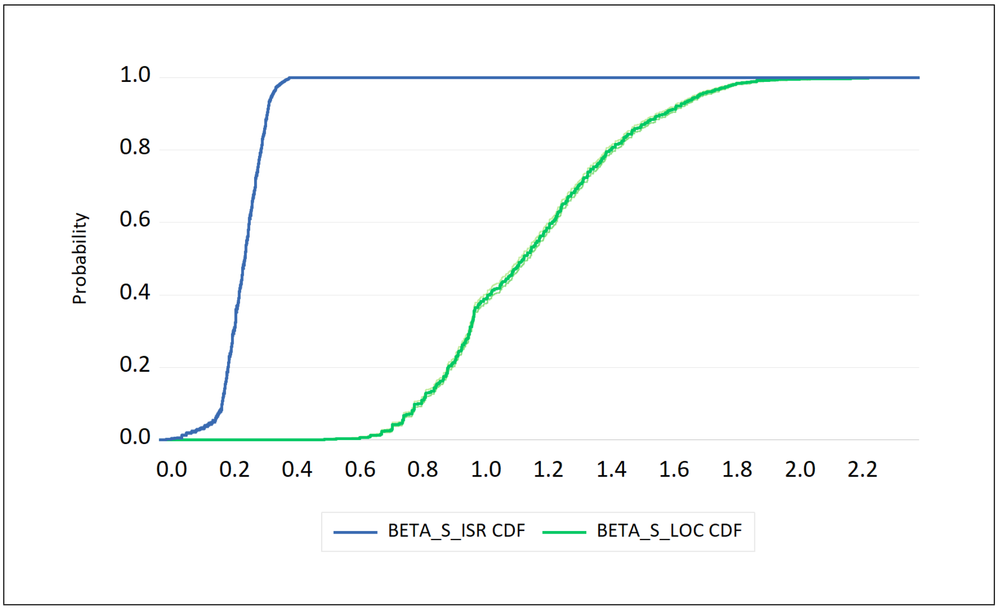
3. The inner fences are defined as the first quartile minus 1.5*Inter Quartile Range (IQR) and the third quartile plus 1.5*IQR. The inner fences are not drawn but the whiskers and staples show the values that are outside the first and third quartiles, but within the inner fences.

4. The staple is a line drawn at the last data point within (or equal to) each of the inner fences. Whiskers are lines drawn from each hinge to the corresponding staple.

Figure 5b: Skills Distributions

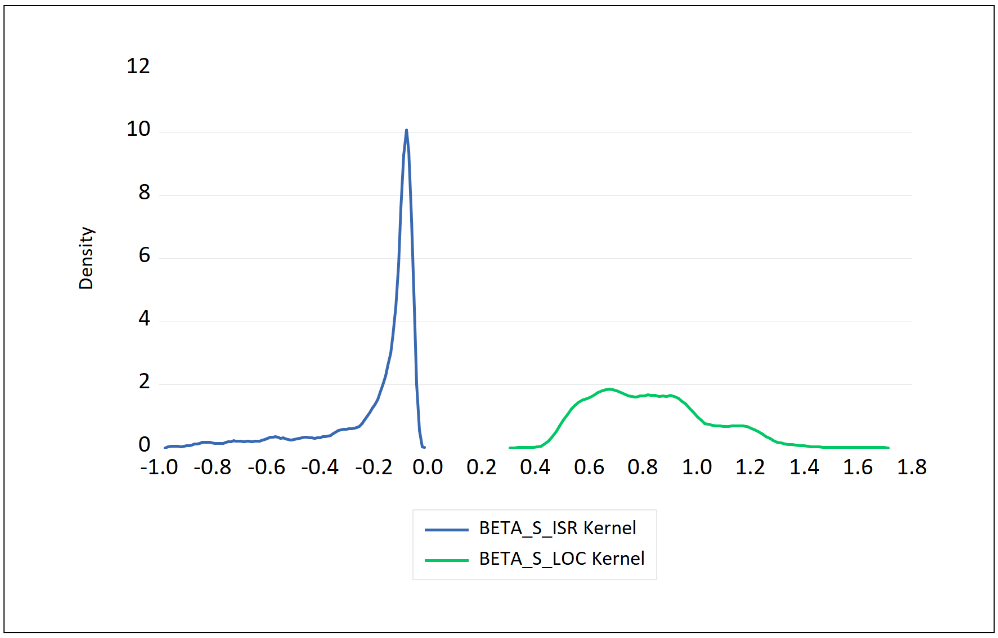
a. 1981

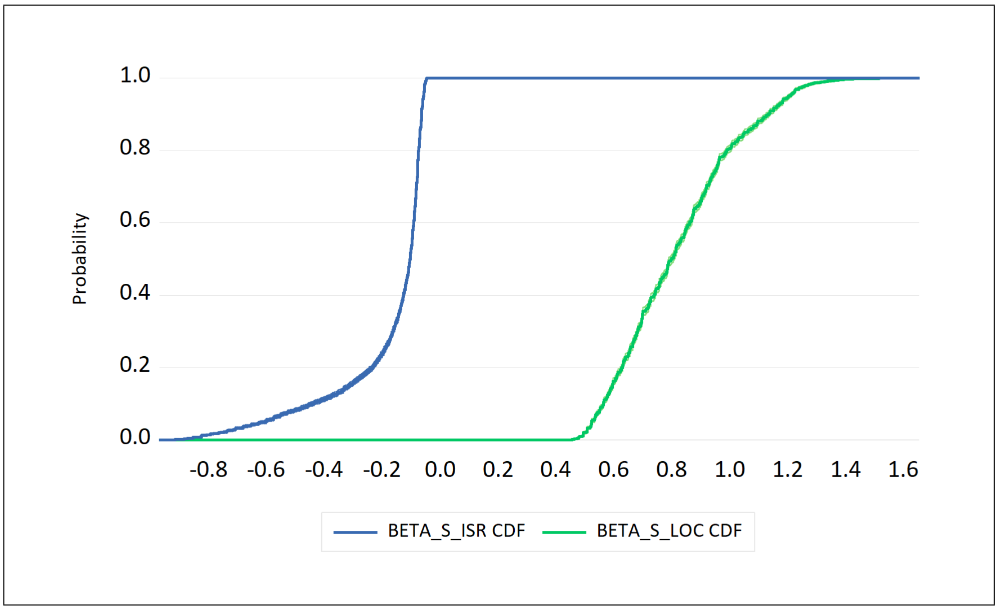




No intersection found. Local dominates Israel.

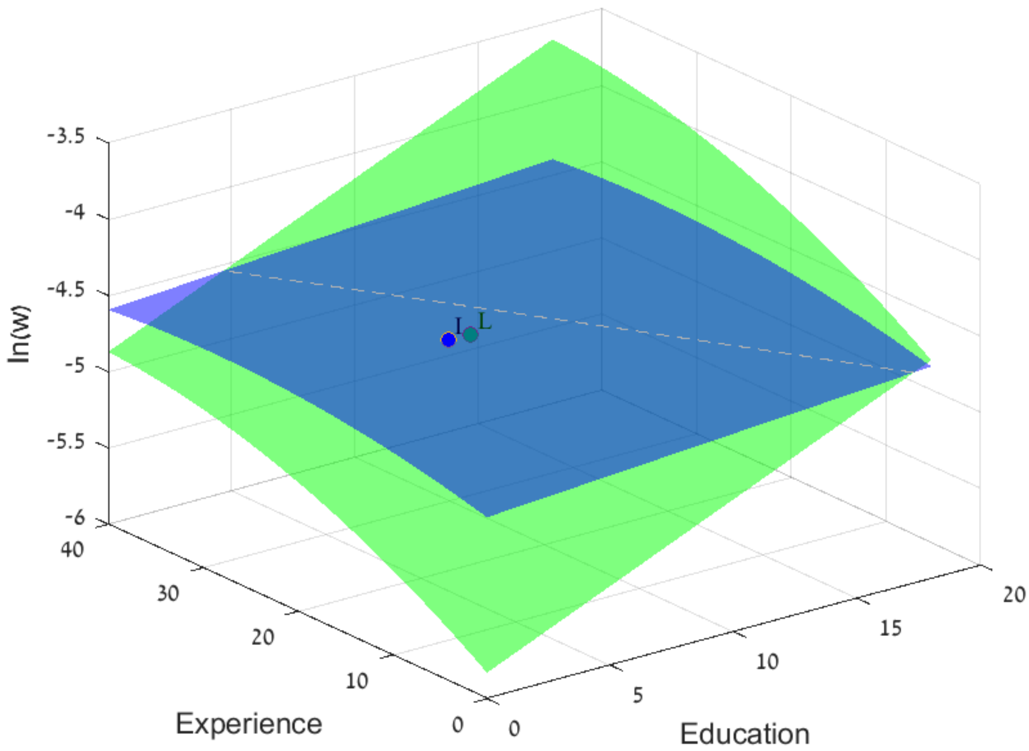
b. 1987



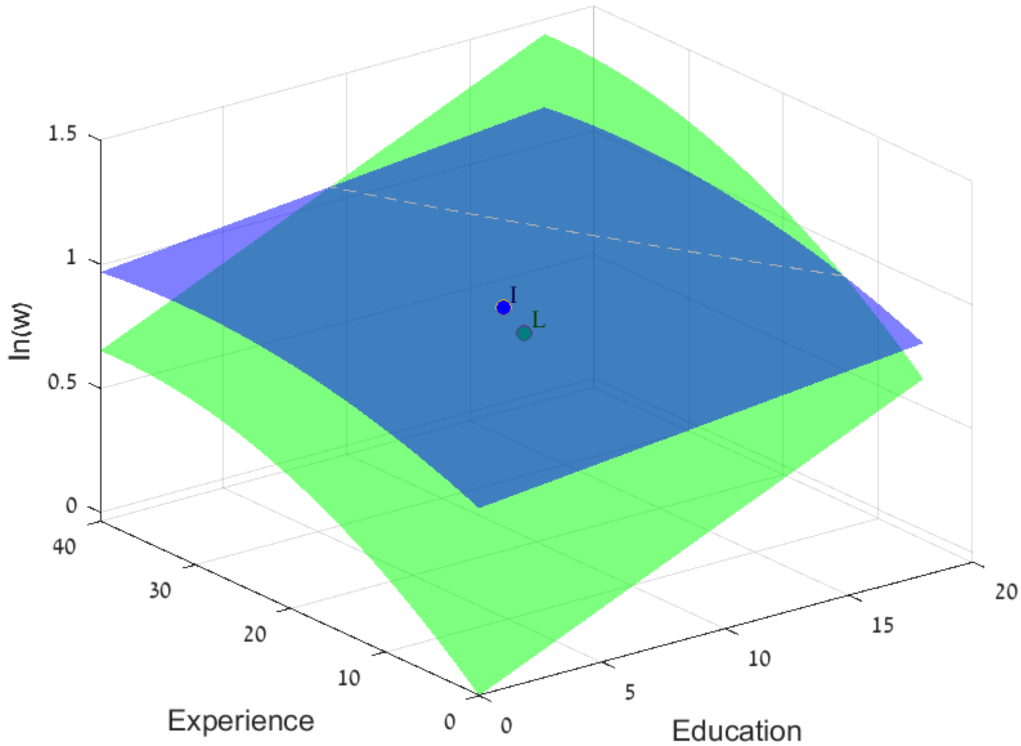


No intersection found. Local dominates Israel.

Figure 5c: Wages as a Function of Skills



a. 1981 log wage equation



b. 1987 log wage equation

Notes:

a. The graphs depict the equation

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

where $S_{\text{exp},i}$ is experience, where

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

$Ev_i = 0$.

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 the table below for workers in the local and Israeli economy, respectively.

Skills and Log Wages Across Locations

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

Notes:

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

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

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

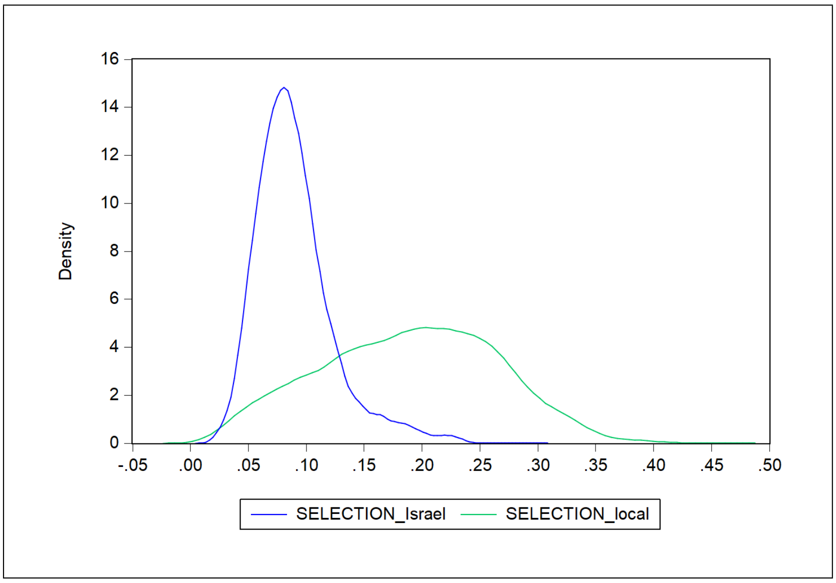
where $S_{\exp,i}$ is experience, where

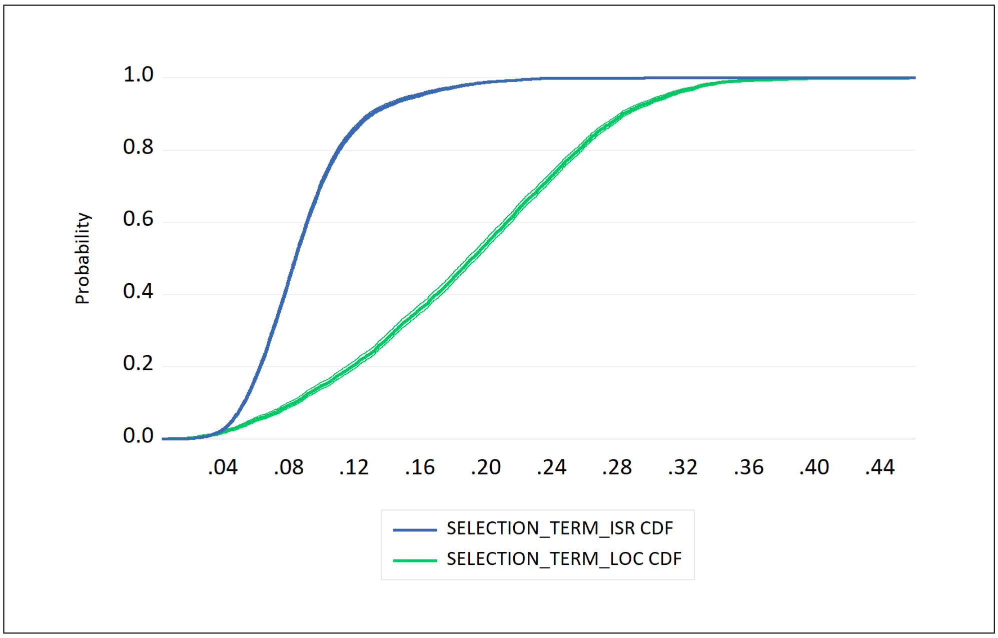
$$C_i = \widehat{k}_i + \beta_{i,0} + E\left(\widehat{\rho}_i \widehat{\sqrt{\sigma_i}} \widehat{\lambda}_i \right) + Ev_i$$

$Ev_i = 0$.

Parameter estimates are taken from Table 3.

Figure 6
a. 1981





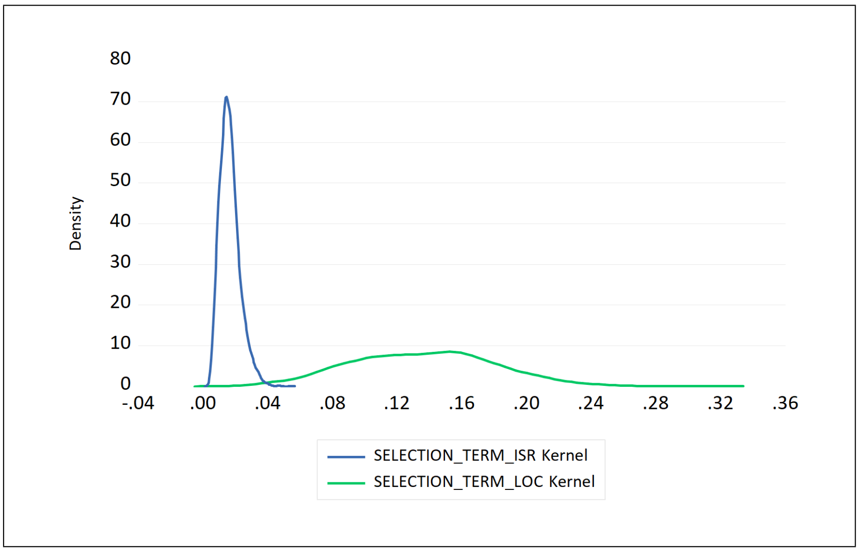
Intersection Number	Intersection Point			Case
	Exact	Min. range	Max. range	
1	.	0.030	0.030	B
2	.	0.030	0.030	B
3	∞	.	.	A

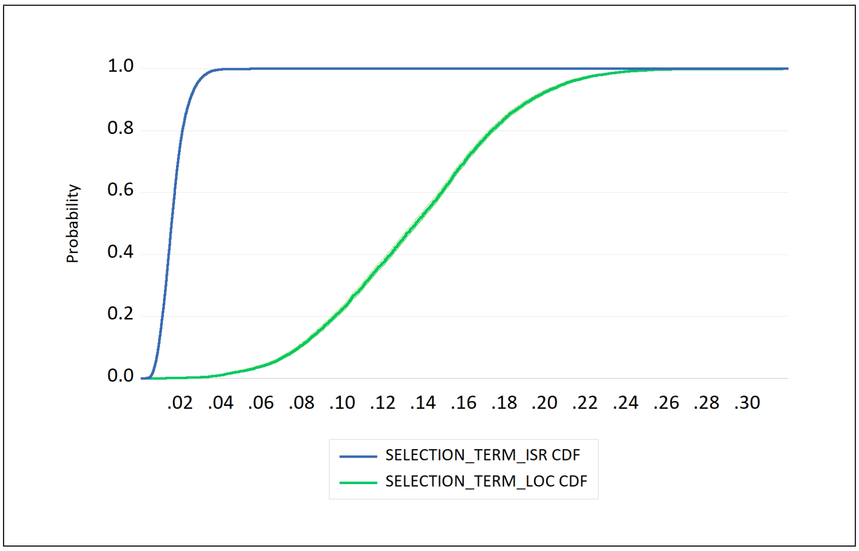
Notes :

Case A: Before this intersection, Local dominates Israel

Case B: Before this intersection, Israel dominates Local.

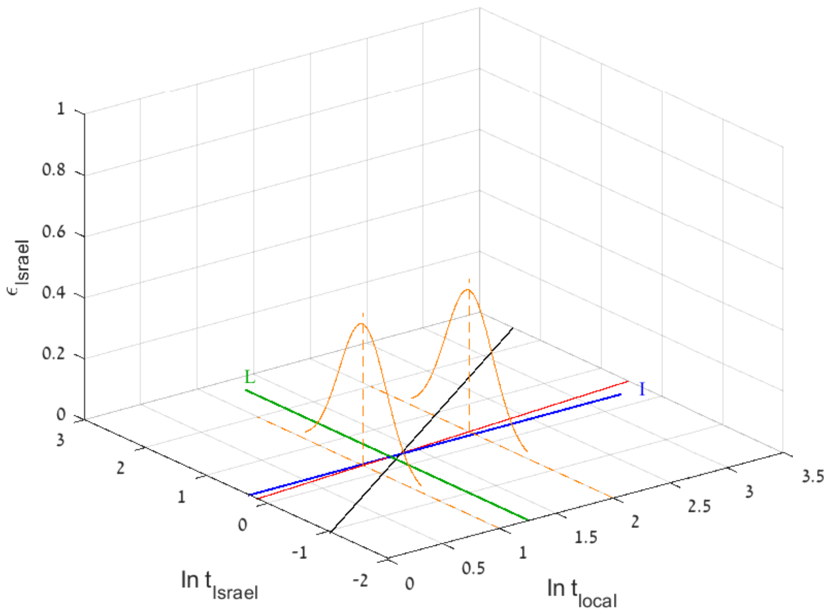
b. 1987



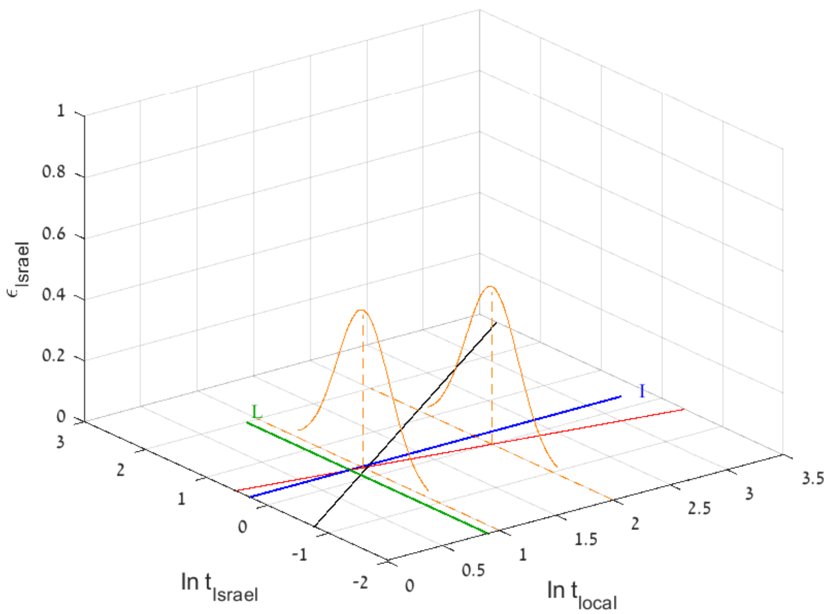


No intersection found. Local dominates Israel.

Figure 7: Tasks and Unobserved Skills



a. 1981 estimates



b. 1987 estimates

Notes:

1. The red regression line expresses equation (29) and 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 2.