HOV Redux:

Identification, HOV, and Factor Biased Technology*

Peter M. Morrow[†] Daniel Trefler [‡]

5 July 2015

ABSTRACT:

The Heckscher-Ohlin-Vanek (HOV) model lies in limbo; every victory offset by criticisms both minor and major. This paper aims to provide some resolution. Using a system of equations approach, we find that a productivity adjusted FPE (PFPE) version of HOV works far better than expected: there exist factor augmenting productivity terms that allow the model to be consistent with approximately 90% of the variation in each of three equations describing the factor content of trade, relative factor use, and wages. Differences in previous results are due to seemingly inconsequential differences in identification strategies. When we allow for a failure of PFPE, we find only minor improvements. The same holds for non-traded services. We also highlight issues of the separate identification of the importance of departures from PFPE versus factor augmenting productivity and relate this to the existing literature.

*We especially thank Susan Zhu for help with the data used in this paper. We also thank Matilde Bombardini, Dave Donaldson, Keith Head, Angelo Melino, Andrés Rodríguez-Clare and seminar participants at Colorado, Essex, The London School of Economics, Oxford, Simon Fraser, Toronto, UBC, UC Berkeley, UC Santa Cruz, and the West Coast Trade Conference (Stanford) for helpful comments and suggestions. Peter Morrow thanks the Economics Department at the University of California, Berkeley where he was a Visiting Scholar during the writing of this paper for their outstanding hospitality.

†Department of Economics, University of Toronto, (peter.morrow@utoronto.ca).

‡Rotman School of Management, University of Toronto, CIFAR, NBER (dtrefler@rotman.utoronto.ca).

1. Introduction

The notion of relative factor abundance has largely disappeared from the empirical international trade literature. A major reason is that the primary model that trade economists use to test its importance, Heckscher-Ohlin-Vanek (HOV), has been in a state of limbo for well over a decade. For every success such as Trefler (1993), there are criticisms that range from mild (Nishioka, 2012, Puzzello, 2012) to completely withering (Gabaix, 1997a, Davis and Weinstein, 2001).[1] The HOV theorem and its minor supply-side modifications now lie like a patient etherized upon a table.

This state of affairs is surprising given the voluminous evidence that relative factor abundance matters for a multitude of economic outcomes. Policy makers continue to base policies on the 'fact' that labor-abundant countries are locked in the Rybcyznski poverty trap of producing labor-intensive products [e.g. Cai, Harrison, and Lin (2014)]; Romalis (2004) shows that differences in relative factor abundance are essential for thinking about industrial specialization across countries and time; Burstein and Vogel (2012) use relative factor abundance as a partial explanation for differences in factor prices both within and across countries. For this reason, we revisit HOV in hope of understanding its past successes and failures, and its relevance for understanding how factor prices and productivity correlate with the process of economic development.

As a foundation for this analysis, we start by showing that the entire empirical content of HOV can be represented as a system of three equations for each factor of production. First, a trade equation describes the factor content of trade. Second, a technology equation describes international differences in average direct unit input requirements. Last, a wage equation imposes factor market clearing consistent with full employment. This system of equations allows us to test the universe of supply-side moments that have previously been considered in the vast HOV literature.

We first examine a special case of our model in which productivity adjusted FPE (PFPE) holds as in Trefler (1993) and Gabaix (1997b). Using data on labor and human capital for 40 countries, we calibrate factor-augmenting international productivity differences that fit our system of equations *superbly*.[2] We are able to explain approximately 90% of the variation of *each*

[1]Other criticisms come from Trefler and Zhu (2000), Fadinger (2011), Maskus and Nishioka (2009) and Marshall (2012). We would be remiss if we did not mention recent work that works to integrate issues of factor abundance into modern models such as the Eaton and Kortum (2002) framework. These include Bernard, Redding and Schott (2007), Parro (2013), Costinot and Rodríguez-Clare (2014), Levchenko and Zhang (2014).

[2]We do not examine physical capital because, as Fitzgerald and Hallak (2004) show, capital is generally mobile across countries and is essentially a traded good, not a factor.

of the trade, technology, and wage equations. Importantly, the calibrated productivity parameters that generate this fit provide good descriptions of not just the observed factor content of trade, but also relative factor use and wages, the latter two of which are less reliant on some of the demand side assumptions that underlie the earlier. Because they fit data on each of these relationships, they fit the full empirical content of the HOV model. The superb fit holds when examining the data factor by factor, as has been the sole convention in the literature, and also when examining data that has been differenced across factors so as to examine the role of *relative* endowments.

Our system of equations approach nests previous 'single equation' approaches to assessing HOV. Trefler (1993) assumes that PFPE holds and uses the trade equation to calculate differences in factor augmenting productivity. Gabaix (1997b) maintains the assumption of PFPE but calculates productivity differences from the wage equation. Remarkably, each comes to a different conclusion regarding the empirical validity of the PFPE version of the HOV model despite using the same model, assumptions, and data. Davis and Weinstein (2001) offer a third single equation approach that uses data on input techniques to assess technology differences. While they find mild support for their version of HOV, they ultimately conclude that non-traded goods and non-zero costs of trade are necessary to explain the data.[3] In sum, each of the single equation approaches offers conflicting evidence for the HOV model.

We show that the differences between our results and each of these single-equation approaches are due to seemingly inconsequential differences in calibration strategies. Specifically, while each of the single equation methods offers very similar values of factor augmenting productivity, they offer very different predictions for the factor content of trade. We show that this is due to the presence of missing trade relative to the predictions of Vanek (1968), which causes the mapping from productivity to the factor content of trade to be very sensitive. We also show that the Vanek equation (our trade equation) can be interpreted an open economy extension of models commonly found in development accounting [e.g. Caselli (2005), Caselli and Coleman (2006)] providing a tight link between these two literatures and showing that there is nothing 'special' or 'rigged' about productivity as measured by the Vanek equation.

There is an important exception to our superb fit: the productivity-adjusted wage of labor *relative* to human capital displays significant departures from PFPE as suggested by Davis and

[3]Paraphrasing Woody Allen, they famously state that the profession continues to rely on the Hecksher-Ohlin model not only because it is so intuitively plausible but also "because we need the eggs."

Weinstein (2001). Accordingly, our second set of results examines the case where PFPE fails. We do this by introducing Ricardian elements that allow wages to be partially determined by factor augmenting productivity as in Trefler (1993) but also by relative endowments as in Davis and Weinstein (2001). It is also in line with the recent literature in macro-development on the importance of agricultural productivity and development, as well as research on international quality differentials since quality and productivity tend to be isomorphic in these models.[4] Although allowing PFPE to fail improves the R^2 of each model, it only does so by a few percentage points suggesting that departures from PFPE are of only minor importance for HOV.

We also show that, when PFPE fails, one can only identify reduced form technology parameters that describe the joint effect of differences in factor prices, factor augmenting productivity, and the elasticity of substitution in production. Following the insights of Diamond, McFadden and Rodriguez (1978), we show that this non-identification is a generic feature of multi-factor models that feature departures from PFPE, factor augmenting productivity differences, and a non-unitary elasticity of substitution. We then show how to extract factor augmenting productivity from these reduced form technology parameters.

To derive our system of equations, we focus on the supply side of the model. In doing so, we make a number of assumptions on the demand side including homothetic preferences and costless trade. As Trefler and Zhu (2010) show, these are both necessary and sufficient for the existence of a Vanek equation. While these assumptions may seem controversial, they yield important insights including the identification issues first formalized by Diamond et al. (1978). This is not to say that demand side considerations do not matter at all.[5] To examine what additional insights can be garnered from the demand side, we follow Davis and Weinstein (2001) and examine the non-traded nature of one particular service industry: government services. We find that this consideration improves the R^2 of each equation by 4-5 percentage points suggesting only minor improvements once one has introduced factor-augmenting technology differences: our three equations fit extremely well without 'fixing' these demand-side departures and fixing them offers only minor improvements. This also suggests that adding trade costs, which we do not consider, will only offer minor improvements to the goodness of fit.

[4]For work on the importance of agricultural productivity see Restuccia, Yang and Zhu (2008), Lagakos and Waugh (2013), Gollin, Lagakos and Waugh (2014). For important work on the importance of international differences in quality see Schott (2004), Hallak (2006), Khandelwal (2010), and Sutton and Trefler (forthcoming).

[5]There is a large and valuable literature on salvaging aspects of the HOV theorem on the demand-side. These include important contributions by Hunter and Markusen (1988) and Justin Caron (2014).

Section 2 presents the theoretical framework that informs our empirical exercise along with the structural system of equations that provide the foundation for our empirical analysis. Section 3 discusses the special case of our model in which PFPE holds. Section 4 discusses the data that we use in our exercise. Sections 5 and 6 present results for the cases where PFPE holds and fails, respectively. Section 7 discusses our calibrated values of factor augmenting productivity. Section 8 concludes.

2. The Model

The primary goals of this section are to derive a system of equations that i) reflects the entire theoretical content of the HOV model, ii) we can use to identify international differences in factor augmenting productivity, and iii) provide over-identifying restrictions that can assess whether our technology terms are consistent with factor market clearing and relative factor use as well as the measured factor content of trade.

2.1. *Preferences, Endowments and Technology*

Let $i,j = 1,\ldots,N$ index countries, let $g = 1,\ldots,G$ index goods, and let $\omega \in \Omega_{gi}$ index varieties of a horizontally differentiated good g produced in country i as in Krugman (1980). The words 'good' and 'industry' are used interchangeably below. Preferences are internationally identical and homothetic with the nested structure:

$$U = \prod_{g=1}^{G} (U_g)^{\eta_g} \quad \text{and} \quad U_g = \left(\sum_{i=1}^{N} \int_{\omega \in \Omega_{gi}} q_{gi}(\omega)^{\frac{\rho_g - 1}{\rho_g}} d\omega \right)^{\frac{\rho_g}{\rho_g - 1}},$$

where $\rho_g > 1$ is the elasticitiy of substitution, $\eta_g > 0$, $\Sigma_g \eta_g = 1$, and $q_{gi}(\omega)$ is a quantity. Let $p_{gi}(\omega)$ be the corresponding price. We assume that trade is costless so that all consumers worldwide face the same price $p_{gi}(\omega)$. Then the price index associated with U_g is $P_g = \left(\sum_{i=1}^{N} \int_{\omega \in \Omega_{gi}} p_{gi}(\omega)^{1-\rho_g} d\omega \right)^{\frac{1}{1-\rho_g}}$.

Let $f = 1,\ldots,K$ index primary factors such as labor. V_{fi} is country i's exogenous endowment of factor f, V_{fgi} is its employment in industry g, w_{fi} is its price, and let $\mathbf{w_i} = (w_1,\ldots,w_K)$. We assume that factors are mobile across industries within a country and immobile across countries.

Turning to technology, a firm uses both primary factors and intermediate inputs of goods $h = 1,\ldots,G$. The production function is Cobb-Douglas in (*a*) an index of primary factors and (*b*)

CES indexes of each of the G intermediate goods. This results in a unit cost function for $\omega \in \Omega_{gi}$ of the form

$$\phi_{gi}(\mathbf{w_i},\mathbf{p}) = [c_{gi}(\mathbf{w_i})]^{\gamma_{g0}} \prod_{h=1}^{G} P_{gh}^{\gamma_{gh}} \tag{1}$$

where

$$P_{gh} = \left(\sum_{j=1}^{N} \int_{\nu \in \Omega_{hj}} \alpha_{gh}\, p_{hj}(\nu)^{1-\rho_h}\, d\nu \right)^{\frac{1}{1-\rho_h}},$$

$\mathbf{p} = \{p_{hj}(\omega) : \omega \in \Omega_{hj}, \forall h,j\}$ is the set of all product prices, $\nu \in \Omega_{hj}$ indexes input varieties, and the γ_{gh} are positive constants with $\Sigma_{h=0}^{G}\gamma_{gh} = 1$. $c_{gi}(\mathbf{w_i})$ is a constant-returns-to-scale unit cost function associated with primary factors. P_{gh} is the unit cost function associated with the CES index of intermediate good h in the production of good g. The α_{gh} are constants that allow for empirically relevant factor intensity asymmetries.

Marginal costs are $\phi_{gi}(\mathbf{w_i},\mathbf{p})$. Per variety variable costs are $\phi_{gi}(\mathbf{w_i},\mathbf{p})q_{gi}(\omega)$. As is standard in the literature, we assume that fixed costs are proportional to marginal costs and given by $\phi_{gi}(\mathbf{w_i},\mathbf{p})\overline{\phi}_g$ for some constant $\overline{\phi}_g > 0$. Because we can (and need to!) flexibly track all global production chains, this cost function is much more general than what is typically used in the literature (e.g., Bernard et al. (2007), Burstein and Vogel (2011)).

2.2. *Firm Behavior*

Profits for any variety $\omega \in \Omega_{gi}$ are

$$\Pi_{gi} = [p_{gi} - \phi_{gi}(\mathbf{w_i},\mathbf{p})]\, q_{gi} - \phi_{gi}(\mathbf{w_i},\mathbf{p})\overline{\phi}_g.$$

There are two sources of demand for ω: (1) Consumers in country j demand ${p_{gi}}^{-\rho_g} P_g^{\rho_g - 1}\eta_g Y_j$ where Y_j is national income, (2) Producers of variety $\nu \in \Omega_{hj}$ demand $b_{ij}(g,h)\left[q_{hj} + \overline{\phi}_h\right]$ where, by Shephard's Lemma,

$$b_{ij}(g,h;\mathbf{w_j},\mathbf{p}) = \frac{\partial \phi_{hj}(\mathbf{w_j},\mathbf{p})}{\partial p_{gi}}.$$

$b_{ij}(g,h;\mathbf{w_j},\mathbf{p})$ is necessarily complicated notation because we need to track the entire global supply chain. Aggregating over all final and intermediate-input demands for a typical variety $\omega \in \Omega_{gi}$, it is straightforward to show that the firm's demand is $q_{gi} = p_{gi}^{-\rho_g} \kappa_g$ for some constant κ_g.[6] This is formalized in lemma 1.

Lemma 1 $q_{gi} = p_{gi}^{-\rho_g} \kappa_g$ *for some constant* κ_g.

Proof See Appendix A ■

Profit maximization for $\omega \in \Omega_{gi}$ yields the standard optimal price:

$$p_{gi} = \frac{\rho_g}{\rho_g - 1} \phi_{gi}(\mathbf{w_i},\mathbf{p}). \tag{2}$$

Zero profits for $\omega \in \Omega_{gi}$, together with this pricing rule, yields:

$$q_{gi} = (\rho_g - 1)\overline{\phi}_g, \tag{3}$$

such that output per variety is pinned down across producer countries within an industry.

Turning to factor markets, consider the demand for factor f by firm $\omega \in \Omega_{gi}$. By Shephard's Lemma this demand per unit of output is

$$d_{fgi}(\mathbf{w_i},\mathbf{p}) = \frac{\partial \phi_{gi}(\mathbf{w_i},\mathbf{p})}{\partial w_{fi}}.$$

Factor market clearing in country i is thus

$$V_{fi} = \sum_{g=1}^{G} \int_{\omega \in \Omega_{gi}} d_{fgi}(\mathbf{w_i},\mathbf{p}) \left[q_{gi} + \overline{\phi}_g\right] d\omega \ = \sum_{g=1}^{G} n_{gi} d_{fgi}(\mathbf{w_i},\mathbf{p}) \left[q_{gi} + \overline{\phi}_g\right] \tag{4}$$

where $n_{gi} = \int_{\omega \in \Omega_{gi}} d\omega$ is the measure of symmetric firms producing varieties of g in country i.

2.3. *Equilibrium*

Define $\mathbf{n}^* = \left\{n^*_{gi}\right\}_{\forall g,i}$, $\mathbf{w}^* = \left\{w^*_{fi}\right\}_{\forall f,i}$, and $\mathbf{p}^* = \{p^*_{gi}(\omega) \colon \omega \in \Omega_{gi},\ \forall g,i\}$. An equilibrium is a triplet $(\mathbf{w}^*,\mathbf{p}^*,\mathbf{n}^*)$ such that when consumers maximize utility and firms maximize profits, product markets clear internationally for each variety, factor markets clear nationally for each factor, and profits are zero. Market clearing and zero profits imply that all income is factor income ($Y_i = \Sigma_f w_{fi} V_{fi}$) and that trade is balanced. It follows from this definition of equilibrium

[6]There is a minor technical point. If a firm buys from itself then we must keep track of the internal price the firm charges itself (the double marginalization problem). To avoid this we assume that a firm does not use its own output as an input. This requires a slight modification of P_{gh} in equation (1). The α_{gh} must be replaced with $\alpha'_{gh}(\omega,\nu)$ where $\omega \in \Omega_{gi}$ is the input variety and $\nu \in \Omega_{hj}$ is the output variety. Then $\alpha'_{gh}(w,\nu) = 0$ for $g = h$ and $\omega = \nu$ (a firm's purchases of its own output) and $\alpha'_{gh}(\omega,\nu) = \alpha_{gh}$ otherwise. The modified P_{gh}, call it $P'_{gh}(\omega)$, depends on ω, but $P_{gh} = P'_{gh}(\omega)$ because the measure of inputs that a firm buys from a single supplier is 0.

that $(\mathbf{w}^*, \mathbf{p}^*, \mathbf{n}^*)$ is an equilibrium if it satisfies the set of equations (2)–(4). From equation (3), output $q_{gi} = q_g$ is independent of i. Since $q_g = p_{gi}^{-\rho_g}\kappa_g$, it follows that price $p_{gi} = p_g$ is also independent of i.

2.4. *The Factor Content of Trade*

Most previous HOV research has not adequately defined the factor content of trade for the case of traded intermediate inputs and international technology differences.[7] We begin by moving from the variety level to the national level. Let $Q_{gi} = n_{gi}q_g$ be output of g in country i. Note that all industry-level output adjustements (Rybczynski effects) are at the extensive margin because output per variety is pinned down. Let $C_{gij} \equiv n_{gj}p_{gj}^{-\rho_g}P_g^{\rho_g-1}\eta_g Y_i$ be country i's final demand of good g produced in country j. Let M_{gij} be country i's imports of good g from country j, including imports of both final goods and intermediate inputs. Let $X_{gj} = \Sigma_i M_{gij}$ be country j's exports of good g. Let $\mathbf{Q_i}$, $\mathbf{C_{ij}}$, $\mathbf{M_{ij}}$, and $\mathbf{X_i}$ be $G \times 1$ vectors with gth elements of Q_{gi}, C_{gij}, M_{gij}, and X_{gi}, respectively. Let $\mathbf{B_{ij}}(\mathbf{w_i}, \mathbf{p})$ be a $G \times G$ matrix whose (g,h)-th element is $n_{gi}b_{ij}(g,h)$. This is what a typical producer of a variety of good h in country j demands (per unit of output) from all n_{gi} producers of good g in country i. Define the matrices:

$$\mathbf{Q} \equiv \begin{bmatrix} \mathbf{Q_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{Q_N} \end{bmatrix}, \qquad \mathbf{C} \equiv \begin{bmatrix} \mathbf{C_{11}} & \cdots & \mathbf{C_{N1}} \\ \vdots & \ddots & \vdots \\ \mathbf{C_{1N}} & \cdots & \mathbf{C_{NN}} \end{bmatrix},$$

$$\mathbf{T} \equiv \begin{bmatrix} \mathbf{X_1} & -\mathbf{M_{21}} & \cdots & -\mathbf{M_{N1}} \\ -\mathbf{M_{12}} & \mathbf{X_2} & \cdots & -\mathbf{M_{N2}} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{M_{1N}} & -\mathbf{M_{2N}} & \cdots & \mathbf{X_N} \end{bmatrix} \text{ and } \mathbf{B} \equiv \begin{bmatrix} \mathbf{B_{11}} & \mathbf{B_{12}} & \cdots & \mathbf{B_{1N}} \\ \mathbf{B_{21}} & \mathbf{B_{22}} & \cdots & \mathbf{B_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B_{N1}} & \mathbf{B_{N2}} & \cdots & \mathbf{B_{NN}} \end{bmatrix}$$

where **Q**, **C**, and **T** are $NGxN$ and **B** is $NGxNG$. The off-diagonal sub-matrices of **B** track trade in intermediate inputs. Output is used either for intermediates (**BQ**) or consumption final demand (**C**) or trade **T** so that $\mathbf{Q} = \mathbf{BQ} + \mathbf{C} + \mathbf{T}$ or

$$\mathbf{T} = (\mathbf{I_{NG}} - \mathbf{B})\mathbf{Q} - \mathbf{C} \tag{5}$$

where $\mathbf{I_{NG}}$ is the $NG \times NG$ identity matrix.

[7]For example, Davis and Weinstein (2001) assume that there are no traded intermediates and Antweiler and Trefler (2002) use a definition of the factor content of trade that is only meaningful if all the assumptions of the HOV model hold.

Let $\mathbf{D_{fi}}(\mathbf{w_i},\mathbf{p})$ be a $1 \times G$ vector with gth element $d_{fgi}(\mathbf{w_i},\mathbf{p})$. Define the $1 \times NG$ vectors $\mathbf{D_f} \equiv [\mathbf{D_{f1}} \cdots \mathbf{D_{fN}}]$ and $\mathbf{A_f} \equiv \mathbf{D_f}(\mathbf{I_{NG}} - \mathbf{B})^{-1}$. Let $\mathbf{T_i}$ be the ith column of $\mathbf{T}$. Then, as shown in Trefler and Zhu (2010, Theorem 1), $F_{fi} = \mathbf{A_f}\mathbf{T_i}$ is the amount of factor f employed worldwide to produce country i's net trade vector $\mathbf{T_i}$.[8] F_{fi} is also the HOV-relevant definition of the factor content of trade for a world with traded intermediate inputs and non-identical technologies.

Theorem 1 *Trefler and Zhu (2010) Let $F_{fi} \equiv \mathbf{A_f}\mathbf{T_i}$. Then*

$$F_{fi} = V_{fi} - s_i \sum_{j=1}^{N} V_{fj}$$

where s_i is country i's share of world income.

Proof See Appendix A. ■

3. Empirical Specification with PFPE

In this section we develop our three equations of empirical interest, namely, the wage, HOV, and technology equations. Following Trefler (1993), we make two assumptions. First, all international differences in technology are factor augmenting:

$$c_{gi}(w_{1i}, \ldots, w_{Ki}) = c_g(w_{1i}/\pi_{1i}, \ldots, w_{Ki}/\pi_{Ki})$$

for all i where the π_{fi} are factor augmenting productivity parameters and the w_{fi}/π_{fi} are factor prices in productivity adjusted units. We choose the normalization $\pi_{f,US} = 1 \; \forall f$, but continue to carry the $\pi_{f,US}$ terms throughout. Second, productivity adjusted factor prices are equalized. That is, $w_{fi}/\pi_{fi} = w_{f,US}/\pi_{f,US}$ for all i or

$$\frac{w_{fi}}{w_{f,US}} = \frac{\pi_{fi}}{\pi_{f,US}}. \qquad (W)$$

'W' is for Wage.

Let $\mathbf{D_f^*}$ be the same as $\mathbf{D_f}$ defined above, but with typical element $d_{fg,US}(\mathbf{w_{US}},\mathbf{p})$ in place of $d_{fgi}(\mathbf{w_i},\mathbf{p})$. Then $F_{fi}^* \equiv \mathbf{D_f^*}(\mathbf{I_{NG}} - \mathbf{B})^{-1}\mathbf{T_i}$ is the factor content of trade in productivity-adjusted (U.S.) units. Under PFPE, the Vanek takes the form described in the following lemma.

[8] Trefler and Zhu (2010) explain this as follows. "Q is referred to as 'gross' output and $C+T$ as 'net' output (or final demand). Let Z_i be an arbitrary net output vector for some country i. Z_i is $NG \times 1$, reflecting the fact that we must keep track not only of what intermediate inputs were used to produce Z_i, but also of where these inputs came from. BZ_i is the vector of intermediate inputs directly needed to produce Z_i. Further, $B(BZ_i) = B^2 Z_i$ is the intermediate inputs directly needed to produce BZ_i. Less abstractly, a sports car consumed in country i (an element of Z_i) requires steel (an element of BZ_i) which requires iron (an element of $B^2 Z_i$) and so on. Thus, $\Sigma_{n=1}^{\infty} B^n Z_i$ is the matrix of intermediate inputs directly and indirectly needed to produce Z_i. Turning from intermediate requirements to total requirements, delivering net output Z_i requires gross output of $Z_i + \Sigma_{n=1}^{\infty} B^n Z_i = \Sigma_{n=0}^{\infty} B^n Z_i = (I_{NG} - B)^{-1} Z_i$." $A_f Z_i$ converts this gross output into factor employment.

Lemma 2

$$F^*_{fi} = \pi_{fi}V_{fi} - s_i \sum_{j=1}^{N} \pi_{fj}V_{fj} \ , \qquad (V)$$

Proof See Appendix A ■

where 'V' is for Vanek. Recalling that $d_{fgi}(\mathbf{w_i},\mathbf{p}) = \partial\phi_{gi}(\mathbf{w_i},\mathbf{p})/\partial w_{fi}$, lemma 2 also shows that our two assumptions imply $\pi_{fi}d_{fgi}(\mathbf{w_i},\mathbf{p}) = \pi_{f,US}d_{fg,US}(\mathbf{w_{US}},\mathbf{p})$ or $d_{fg,US}/d_{fgi} = \pi_{fi}/\pi_{f,US}$. We aggregate this up to the factor level (which is the level of the wage and HOV equations) by taking weighted averages:

$$\overline{d}^{-1}_{fi} \equiv \sum_{g=1}^{G} \theta_{fgi}\frac{d_{fg,US}}{d_{fgi}} = \frac{\pi_{fi}}{\pi_{f,US}}, \qquad (T)$$

where $\theta_{fgi} \equiv V_{fgi}/V_{fi}$ is the share of V_{fi} that is employed in industry g.[9] The θ_{fgi} are data. T is for Technology. Equations (W), (V), and (T) are our three equations of interest.

4. The Data

The database contains 40 developed and developing countries and 24 ISIC industries. Countries and industries are listed in Appendix B (the Data Appendix). All data are for 1997. There are two factors (labor L and human capital H). With the exception of human capital, data are similar to those used in Trefler and Zhu (2010). Trade data (T_i), input-output data ($b_{ij}(g,h)$) and output data (Q_{gi}) are from GTAP. We discuss the construction of **B** in Appendix B. Data on labor by industry (V_{Lgi}) are from multiple sources. For OECD countries, data are from the STAN database. For non-OECD countries, data are from the UNIDO INDSTAT database (manufacturing) and the ILO LABORSTA database (non-manufacturing). Employment by industry is scaled so that country-level employment sums to Barro and Lee (2013) workforce totals.

Wages for labor L are calculated as the average of wages in unskilled occupations. We use unskilled occupations that have low levels of human capital and produce tradable goods e.g., labourers and sewing-machine operators. See Appendix B for details. Data are from the Occupational Wages Around the World (OWW) database (Freeman and Oostendorp (2012), available at `www.nber.org/oww`), a clean version of data originally collected by the ILO. Human capital is defined as the stock of high school equivalent workers and is calculated by dividing the payroll of each industry by the wage of a high school equivalent worker in that country. The wage of a high school equivalent worker is calculated by using the wage of unskilled labor from OWW,

[9] $\Sigma_g\theta_{fgi} = 1$

educational attainment data from Barro and Lee (2013), and estimates of the return to schooling across countries from Banerjee and Duflo (2005). Appendix B details our procedure.

Each observation in our three equations (W, V and T) is subscripted by factor and country (f,i). We scale observation (f,i) by $s_i^{1/2}$. We never pool across factors and so do not need to scale by f. However, to make figures more readable we scale observation (f,i) in the V and V' equations by the standard deviation of $\varepsilon_{fi} \equiv F_{fi} - (V_{fi} - s_i V_{fW})$.[10]

5. Results with PFPE

We begin with a useful preliminary result. For each of our three equations there is a unique vector of productivities that makes the equation fit perfectly. Setting $\pi_{f,US} = 1$ in what follows,

$$\pi_f^W \equiv (w_{f1}/w_{L,US}, \ldots, w_{fN}/w_{L,US})$$

makes the wage equation for factor f fit perfectly,

$$\pi_f^T \equiv (\overline{d}_{f1}^{-1}, \ldots, \overline{d}_{fN}^{-1})$$

makes the technology equation for factor f fit perfectly and, as shown by Gabaix (1997b), the vector π_f^V with ith element

$$\pi_{fi}^V \equiv \frac{s_i/V_{fi}}{s_{US}/V_{f,US}} + \frac{s_i}{V_{fi}} \left(\frac{F_{fi}^*}{s_i} - \frac{F_{f,US}^*}{s_{US}} \right) \tag{6}$$

makes the Vanek equation for factor f fit perfectly.[11]

5.1. Previous 'Single-Equation' Approaches

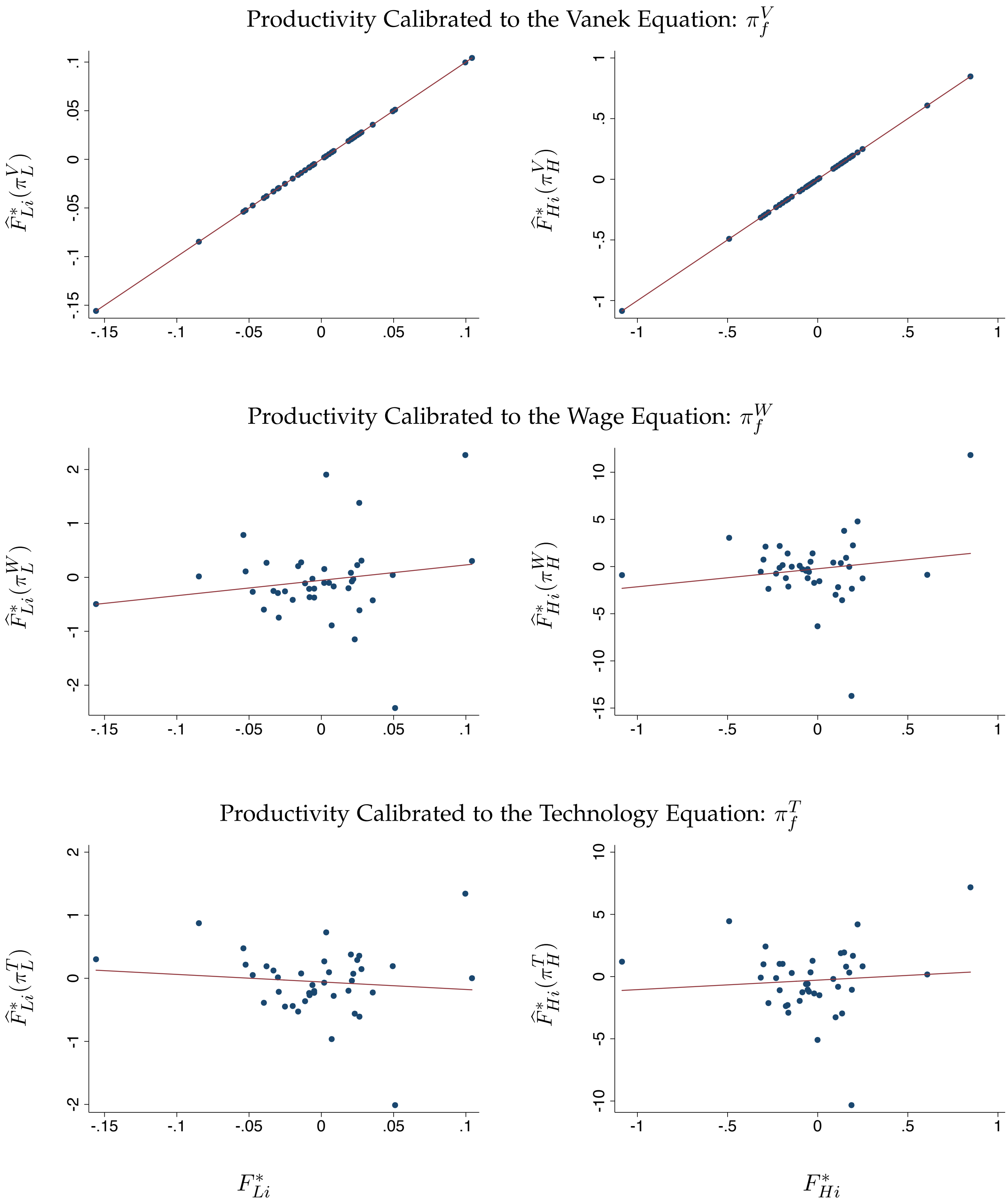
At a high level of abstraction, the entire literature on supply-side modifications of the HOV model can be categorized into one of the following single-equation approaches: (1) Plug π_f^V into the wage equation and evaluate the fit e.g., Trefler (1993). (2) Plug π_f^W into the Vanek equation and evaluate the fit e.g., Gabaix (1997b). (3) Plug π_f^T into the Vanek equation and evaluate the fit e.g., Davis and Weinstein (2001). [12] We refer to these approaches as 'single-equation' because each infers productivity from a single one of the equations above.

[10] That is, by the square root of $\sum_{i=1}^N (\varepsilon_{fi} - \bar{\varepsilon}_f)^2 / (N-1)$ where $\bar{\varepsilon}_f = \sum_{i=1}^N \varepsilon_{fi}/N$.

[11] As we will discuss in 5.5, equation 6 is simply the open economy version of a common expression in development accounting when PFPE holds.

[12] Gabaix (1997b) plugs π_f^W into the Vanek equation and then estimates the following (ignoring trade imbalances): $F_{fi}/s_i = \alpha_f + \beta_f w_{fi} V_{fi}/s_i + \varepsilon_{fi}$. He rejects the Vanek equation because the $\widehat{\beta}_f$ are implausibly smaller than 1.

Figure 1: Performance of the HOV Equation: $\widehat{F}^*_{fi}(\pi_f)$ vs. F^*_{fi}



Notes: Each panel plots the predicted productivity-adjusted factor content of trade $\widehat{F}^*_{fi}(\pi_f)$ from equation (7) against the actual productivity-adjusted factor content of trade F^*_{fi}. The left column is for labour and the right column is for human capital. Each row uses a different π_f and this is indicated in the header to the row. π_f^V makes the Vanek equation fit perfectly (equation V), π_f^W makes the wage equation fit perfectly (equation W) and π_f^T makes the technology equation fit perfectly (equation T).

Turning to single-equation empirics, consider figure 1. Each panel plots F^*_{fi} on the horizontal axis and its prediction

$$\widehat{F}^*_{fi}(\pi_f) = \pi_{fi} V_{fi} - s_i \sum_{j=1}^{N} \pi_{fj} V_{fj} \tag{7}$$

on the vertical axis with π_f obtained in each of the three manners above. An observation is a country. All lines are OLS best fits. The left-hand column is for labor ($f = L$) while the right-hand column is for human capital ($f = H$). In the first row Trefler's procedure is used. That is, F^*_{fi} is plotted against $\widehat{F}^*_{fi}(\pi^V_f)$. By construction, the fit is perfect. In the second row, π^W_f is used, that is, F^*_{fi} is plotted against $\widehat{F}^*_{fi}(\pi^W_f)$. The fit is horrible, which is the same conclusion arrived at by Gabaix (1997b). In the third row, π^T_f is used, that is, F^*_{fi} is plotted against $\widehat{F}^*_{fi}(\pi^T_f)$. Again the fit is horrible, which is the same conclusion arrived at by Davis and Weinstein (2001). In short, we confirm previous findings with our data.

Beyond the poor fit, there is another feature of these panels that will be useful later. Looking down the left-hand L column, the horizontal scale is always the same (it runs from -0.15 to 0.10), but the vertical scale is not. For the top panel it also runs from -0.15 to 0.10, but for the other two panels it runs from -2.00 to 2.00. Similarly for the left-hand H column. This is what Trefler (1995) referred to as the 'Case of the Missing Trade.' Previewing results later in this section, it will prove invaluable for understanding why the prediction $\widehat{F}^*_{fi}(\pi_f)$ is so sensitive to the choice of π_f.

5.2. *Simultaneous 3-Equation Results*

There are two big problems with single-equation approaches. First, they use information from only two of the three equations — valuable information is tossed out. For example, Trefler and Gabaix do not use international data on input requirements and Davis and Weinstein do not use wage data. Second, the loss function is oddly lexicographic. For example, suppose that π_f is calibrated from the wage equation and that a small change in π^W_f greatly improves the fit of the

In Davis and Weinstein (2001), the technology equation is couched in terms of the A_f matrix rather than the D_f matrix. In their H3 hypothesis, they estimate the regression $\ln a_{fgi} = \lambda_i + \lambda_{fg} + \varepsilon_{fgi}$. This is very much like using the technology equation to calibrate π^T_f. They then use the fitted $\ln a_{fgi}$ to construct a fitted factor content of trade $\widehat{F}_{fi} = \widehat{A}_f T_i$ and evaluate the fit of $\widehat{F}_{fi} = V_{fi} - s_i \sum_j V_{fj}$. This is very much like to plugging π^T_f into the Vanek equation and evaluating its fit. We can be more specific while remaining succinct if we ignore intermediates so that $a_{fgi} = d_{fgi}$. Their λ_{fg} plays the role of $\ln d_{fg,US}$ and their λ_i is a Hicks neutral technology difference. Davis and Weinstein do not actuality consider factor augmenting technology differences, which would have required a λ_{fi} in place of λ_i. If they had estimated $\ln d_{fgi}/d_{fg,US} = \lambda_{fi} + \varepsilon_{fgi}$ using weighted OLS with weights $(\theta_{fgi})^{1/2}$ then $e^{-\hat{\lambda}_{fi}} = \pi^T_{fi}$ i.e., they would be exactly using a single-equation approach.

HOV equation without discernibly affecting the fit of the wage equation. Then one would want to consider this small change. Unfortunately, single-equation approaches rule out such a change. We now turn to an alternative approach that uses all three equations and equalizes marginal improvements across equations.

Consider the R^2s associated with each of our equations. Denote these by $R^2_{fk}(\pi_f)$ where $k = V, W, T$ indexes whether it is the Vanek, wage, or technology equation and f indexes factors L and H.[13]

For each f separately, we choose π_f to maximize the minimum of the three R^2s:

$$\widehat{\pi}_f = \underset{\pi_f}{\operatorname{argmax}} \left\{ \min \left\{ R^2_{fV}(\pi_f)\, ,\, R^2_{fW}(\pi_f)\, ,\, R^2_{fT}(\pi_f) \right\} \right\} \ .$$

While there are a number of alternative loss functions we could use, the solution to this maximin loss function has the advantage of having an analytical characterization.

Theorem 2 *(1) $\widehat{\pi}_f$ exists and is unique. (2) $\widehat{\pi}_f$ is a matrix-weighted average of the π_f^V, π_f^W and π_f^T. (3) The two worst-fitting equations have equal R^2s. (4) $\widehat{\pi}_f$ is Pareto optimal: Any π_f that improves the R^2 of one equation lowers the R^2 of at least one of the two remaining equations.*

Proof See Appendix C ■

Part 2 states that $\widehat{\pi}_f$ has an intuitive interpretation as a linear function of the single-equation calibrations.[14] One might expect $\widehat{\pi}_f$ to equate all three R^2s. Part 3 states that, because of possible corner solutions, $\widehat{\pi}_f$ only equates the R^2s of the two worst-fitting equations. Part 4 states that the maximin estimator is decidedly non-lexicographic: It equates marginal benefits across the two worst-fitting equations.[15] To conclude, the maxmin estimator addresses the two problems with single-equation approaches. It uses information from all three equations and efficiently trades off the fit of the two worst-fitting equations.

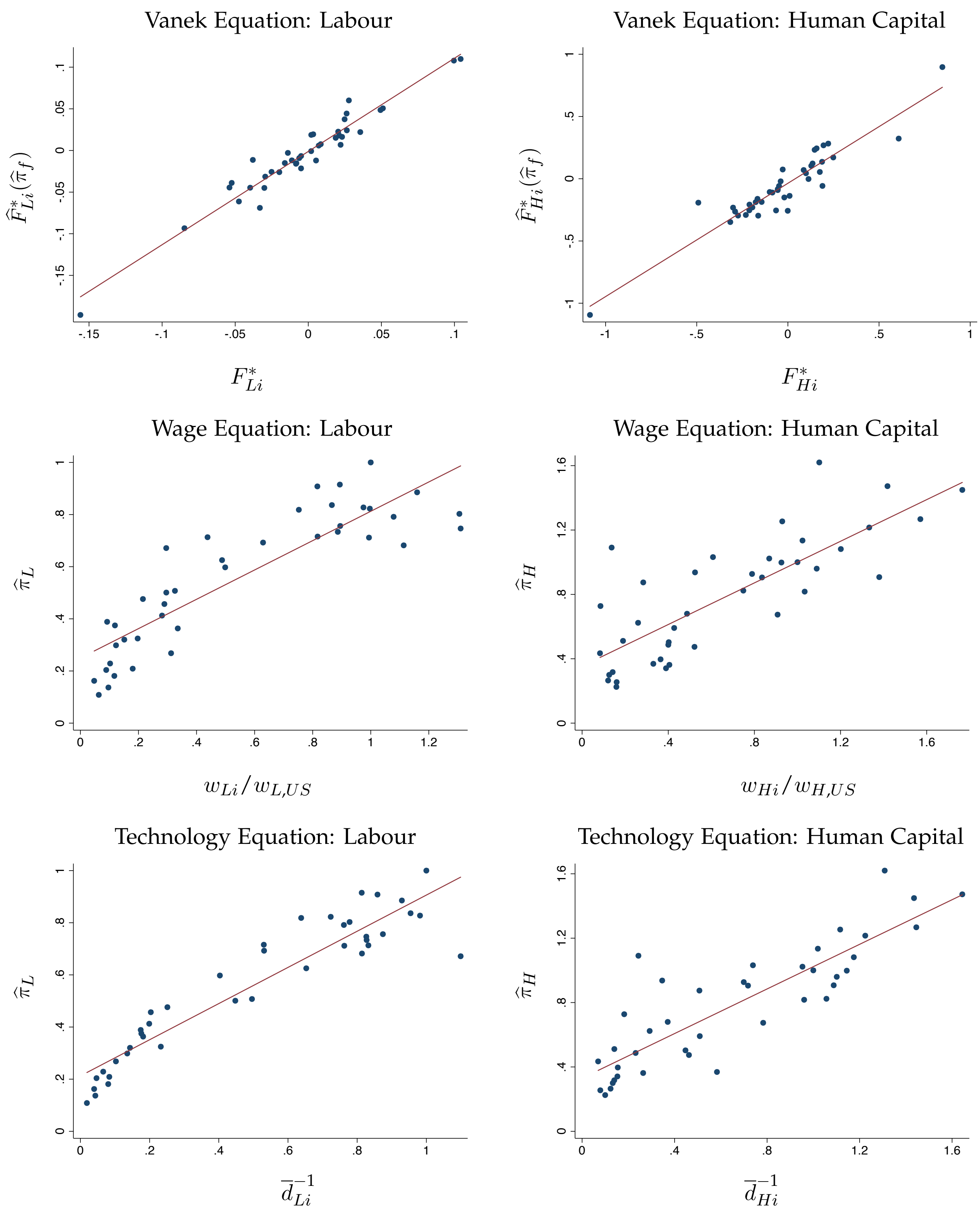
---

[13] E.g., $R^2_{LV}(\pi_L) = 1 - \sum_{i=1}^N \left[F^*_{Li} - \widehat{F}^*_{Li}(\pi_L) \right]^2 / \sum_{i=1}^N \left[F^*_{Li} \right]^2$.

[14] The proof of part 2 is not trivial. A famous result in statistical decision theory is that a minimax or maximin estimator can be represented as a Bayesian estimator for some prior (Ferguson, 1967, ch. 2.9, theorem 1). Further, a linear Bayesian estimator can be represented as a matrix weighted average of the prior and posterior means (Leamer, 1978, theorem 3.9).

[15] It is helpful to think of $R^2_{fk}(\pi_f)$ as a utility function and π_f as a vector off goods. This utility function has an indifference curve through the point $\widehat{\pi}_f$. (Mathematically, the indifference curve is the set of π_f satisfying $R^2_{fk}(\pi_f) = R^2_{fk}(\widehat{\pi}_f)$.) Part 4 states implies that if one looks at the two worst-fitting equations, the two associated indifference curves are tangent at $\widehat{\pi}_f$.

Figure 2: PFPE: Performance of the Vanek, Wage and Technology Equations



Notes: The top row plots the predicted productivity-adjusted factor content of trade $\widehat{F}^*_{fi}(\pi_f)$ from equation (7) against the actual productivity-adjusted factor content of trade F^*_{fi}. The middle row plots factor prices $(w_{fi}/w_{f,US})$ against factor augmenting productivity $(\pi_{fi}/\pi_{f,US})$. The bottom row plots inverse (average) unit input requirements $(d_{fi}/d_{f,US})$ against factor augmenting productivity $(\pi_{fi}/\pi_{f,US})$. The left column is for labour and the right column is for human capital. Each row uses a the same $\hat{\pi}_f$ obtained as discussed in the text.

Table 1: Test Statistics for the Fit of the Vanek Equation

	Labor				Human Capital			
	Rank Corr.	Variance Ratio	Sign Test	Slope Test	Rank Corr.	Variance Ratio	Sign Test	Slope Test
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Three-Equation Approaches								
1. PFPE	0.94	0.746	0.95	0.84	0.90	1.056	0.90	0.96
2. No PFPE	0.97	1.228	0.98	1.08	0.89	1.108	0.95	1.00
3. No PFPE, Non-Traded	0.96	0.998	0.98	0.99	0.98	0.868	0.95	0.92
Single-Equation Approaches								
4. π_f^W from Wage eqn.	0.14	0.004	0.58	0.01	-0.05	0.007	0.55	0.01
5. π_f^T from Tech. eqn.	-0.08	0.007	0.55	-0.01	0.04	0.012	0.60	0.01
6. $\pi_{fi} = 1$	-0.02	0.000	0.38	0.00	-0.21	0.000	0.43	-0.01

Notes: This table presents test statistics for the fit of the Vanek equation (V). Using equation (7), this compares F_{fi}^* to $\widehat{F}_{fi}^*(\pi_f)$. The only difference between rows is that each uses a different π_f. The relevant π_f is identified in the 'Specification' column. 'Rank Corr.' is the rank or Spearman correlation between F_{fi}^* and $\widehat{F}_{fi}^*(\pi_f)$. 'Variance Ratio' is the variance of F_{fi}^* divided by the variance of $\widehat{F}_{fi}^*(\pi_f)$. 'Sign Test' is the proportion of observations for which F_{fi}^* and $\widehat{F}_{fi}^*(\pi_f)$ have the same sign. 'Slope Test' is the OLS estimate of β from the regression $F_{fi}^* = \alpha + \beta \widehat{F}_{fi}^*(\pi_f)$.

Figure 2 presents our calibration results for our maximin estimate $\widehat{\pi}_f$. Each point is a country. The top left panel plots the Vanek equation for labour. The horizontal axis is the actual factor content of trade F_{Li}^*. The vertical axis is the predicted factor content of trade $\widehat{F}_{Li}^*(\widehat{\pi}_L)$ as defined in equation (7). The middle left panel plots the wage equation. The horizontal axis is $w_{Li}/w_{L,US}$ and the vertical axis is $\widehat{\pi}_L$. The bottom left panel plots the technology equation. The horizontal axis is $\overline{d}_{Li}^{-1}$ and the vertical axis is $\widehat{\pi}_L$. The right-hand column plots the corresponding results for human capital.

Figure 2 leads to a very different conclusion about the HOV model with productivity-adjusted factor price equalization. Most previous studies, as well as the single-equation results of figure 1, strongly reject the model. In contrast, figure 2 conveys the impression that the modified HOV model works very well. The labour R^2s are 0.899 for the Vanek equation, 0.936 for the wage equation and 0.899 for the technology equation. The corresponding human-capital R^2s are 0.862, 0.888 and 0.862. Not only does the model perform well, it does so for a wider range of data phenomena than have previously been examined in a single coherent framework: The model simultaneously fits the HOV, wage and technology equations. In plain English, it explains the observed factor content of trade, international differences in factor prices, and international differences in choice of techniques.

Of course, the fit is not perfect, as should be expected given that the $\widehat{\pi}_f$ must simultaneously explain three phenomena. But two things stand out. First, the Vanek plots display three extreme points (countries). These are Hong Kong at the bottom left and Taiwan and Japan at the top right. However, the visual impression conveyed by the Vanek plots does not change at all when these countries are deleted or when these countries are omitted from the plots so that the observations near $(0,0)$ are 'unpacked.' Second, the wage equation for labour displays systematic deviations from PFPE, a point which is developed at length below.

It is conventional in HOV papers to report a large number of test statistics. While we feel that the plots tell the full story, we give a nod to convention in table 1. Row 1 reports statistics corresponding to figure 2. Rows 4 and 5 report statistics corresponding to the second and third rows of figure 1. By construction, the first row of figure 1 has statistics that are all equal to 1.00. Row 6 reports the results for the vanilla Vanek equation ($\pi_{fi} = 1$). The striking contrast between row 1 and rows 4–6 might leave the reader wondering whether we are estimating a data identity or slipping 'gravity' in through the back door. This is simply not the case. Indeed, rows 1, 4 and 5 describe an identical equation, the only difference being the choice of π_f. Yet row 1 performs well and rows 4–5 perform poorly. Thus, the performance differences cannot be explained by data identities or back-door gravity.

5.3. *The Spirit of Heckscher and Ohlin*

When we teach the Heckscher-Ohlin model to our students we focus on the role of *relative* abundance and *relative* factor intensities i.e., human capital relative to labor. Vanek (1968) and Leamer (1980) developed models which are separable across factors so that one need only examine one factor at a time. This has been the tradition in the HOV literature. We thus 'stress test' our results by examining how they perform in *relative* terms.

Figure 3 reports the results after differencing. Remarkably, the Vanek and technology equations do *extremely* well in differences! No previous paper has subjected its results to this kind of

Figure 3: PFPE: Differencing Across Factors

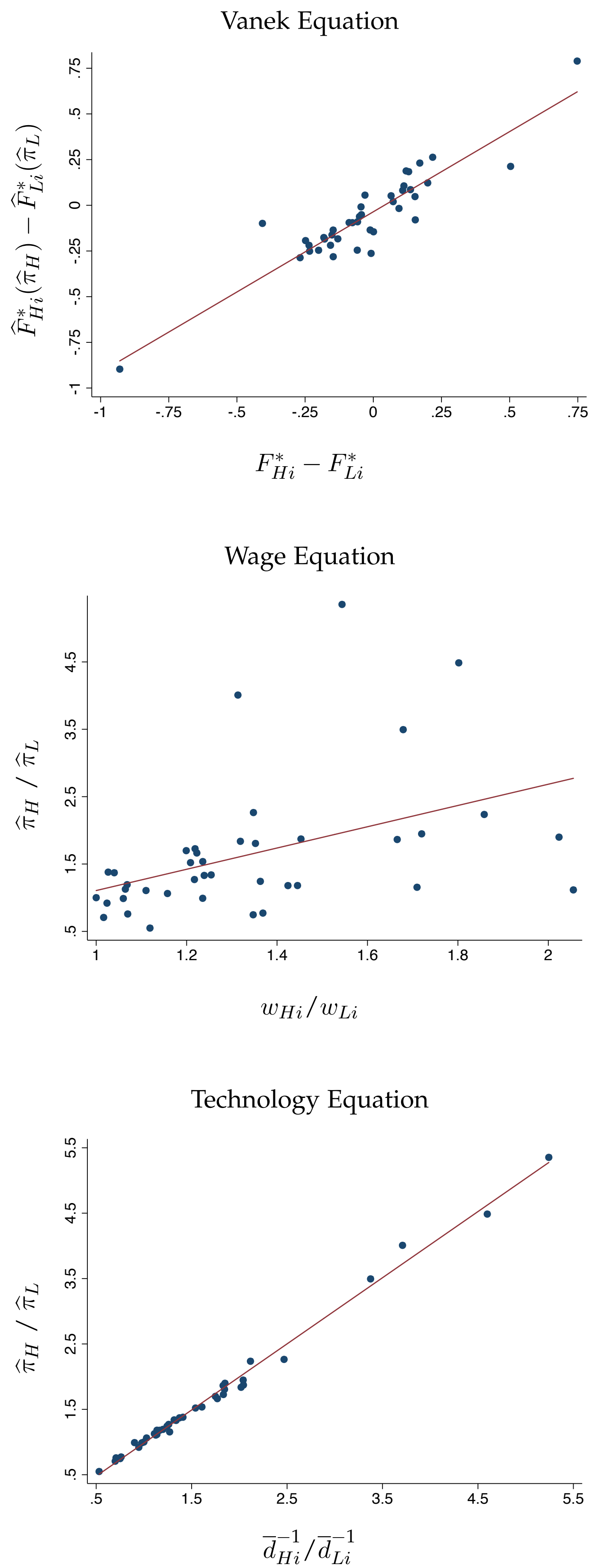


Table 2: π_{fi} Correlation Matrix

	Labor					Human Capital			
	π^V	π^T	π^W	$\hat{\pi}$		π^V	π^T	π^W	$\hat{\pi}$
π^V	1.00				π^V	1.00			
π^T	.94	1.00			π^T	.84	1.00		
π^W	.86	.84	1.00		π^W	.78	.95	1.00	
$\hat{\pi}$	.99	.94	.88	1.00	$\hat{\pi}$	.99	.86	.80	1.00

stress test.[16] Further, these results will be comforting to those interested in the pre-Vanek/Leamer characterization of the Heckscher-Ohlin model. Equally intriguing, the wage equation fits horribly.[17] This is our first clear clue that something is amiss with productivity-adjusted factor price equalization (PFPE). As we shall see, the results for wages improve dramatically when we relax PFPE. But before we go there, we have one last piece of business.

There is another reason this is significant. If the ratio $\widehat{\pi}_H/\widehat{\pi}_L$ equals 1 (or even if it is the same in all countries) then the model is more about Hick's neutral change. The bottom panels show that the ratio is not equal to 1 nor is it equal across countries, and hence Hick's neutral change does not capture the data nearly as well as factor-augmenting technical change.

5.4. *Puzzle #1: Reconciling the Single-Equation and System Results*

There is a truly surprising difference between the negative results in figure 1 and the positive results in figure 2. Recall that in both instances the Vanek equation is of the form $F^*_{fi} = \widehat{F}^*_{fi}(\pi^V_f)$ where F^*_{fi} is data that is the same for all approaches and $\widehat{F}^*_{fi}$ is a function that is the same for all approaches. Hence, there are only two possible explanations for the difference in results. Either π^W_f and π^T_f are very different from π^V_f and $\hat{\pi}_f$ or $\widehat{F}^*_{fi}(\pi_f)$ is a very sensitive function, by which we mean that small changes in its argument leads to large changes in its value.

We can easily dismiss the first possibility. Table 2 presents the correlation between the π^V, π^T, π^W, and $\hat{\pi}$. Clearly, they are very similar. It thus must be that $\widehat{F}^*_{fi}$ is very sensitive to the choice of π_f.

[16]The reader will notice that the trade equation takes the absolute difference whereas the wage and technology equations are in ratios. This is because some small values for $F^*_{fi}(\hat{\pi}_f)$ lead to large outliers that hinder visual inspection.

[17]While the rank correlations for the data in the panels presenting the Vanek and Technology differences are 0.95 and 0.99, respectively, the correlation of fitted relative wages with the actual data is only 0.54.

To show this, we consider two productivity vectors that are close together and show that they yield very different values of $\widehat{F}^*_{fi}$. To reduce notational burden, we take these to be π^W_f and $\widehat{\pi}_f$ and define the (small) difference $\varepsilon_f = \pi^W_f - \widehat{\pi}_f$. Inspection of equation (7) shows that $\widehat{F}^*_{fi}$ is linear in its arguments. Hence $\widehat{F}^*_{fi}(\pi^W_f) - F^*_{fi}(\widehat{\pi}_f) = \widehat{F}^*_{fi}(\varepsilon_f)$ or

$$\widehat{F}^*_{fi}(\pi^W_f) - F^*_{fi}(\widehat{\pi}_f) = \varepsilon_{fi} V_{fi} - s_i \sum_{j=1}^{N} \varepsilon_{fj} V_{fj}$$

where ε_{fi} is the ith element of ε_f. Now consider the variance of the right-hand side. Suppose that the ε_{fc} are purely random variables with mean 0 and small variance $\sigma^2_{\varepsilon f} \approx 0.02$. Then the right-hand side is 0 on average. Its variance is $\sigma^2_{\varepsilon f}\text{var}[V_{fi} - s_i \sum_j V_{fj}]$. Let σ^2_{Ff} and σ^2_{Vf} be the variances of F^*_{fi} and $[V_{fi} - s_i \sum_j V_{fj}]$, respectively, where the variation is across observations i. Because missing trade is so severe, the variance ratio is $\sigma^2_{Ff}/\sigma^2_{Vf} = 0.0002$. Hence the variance of the right-hand side is $\sigma^2_{\varepsilon f}\sigma^2_{Vf} = \sigma^2_{\varepsilon f}\sigma^2_{Ff}/0.002 = 5{,}000\sigma^2_{\varepsilon f}\sigma^2_{Ff}$! Thus, even though $\sigma^2_{\varepsilon f}$ is small, the right-hand side has a large variance relative to the variance of what is to be explained (σ^2_{Ff}). Restated, $\widehat{F}^*_{fi}(\pi^W_f)$ and $F^*_{fi}(\widehat{\pi}_f)$ may be equal on average, but because of missing trade, the former has a much larger variance. Single-equation approaches are like drunk dart players: Every dart completely misses the dartboard, but if you average them you get a bull's-eye.

One can see this graphically by comparing the second row of figure 1 with the first row of figure 2. From the former, $\widehat{F}^*_{fi}(\pi^W_f)$ is centred on 0, but has a range of roughly $(-2{,}2)$. From the latter, $\widehat{F}^*_{fi}(\widehat{\pi}_f)$ is also centred on 0, but has a much smaller range of roughly $(-0.15{,}0.1)$.

This establishes that the function $\widehat{F}^*_{fi}$ is very sensitive to the choice of π_f. It thus explains the discrepancy in results between single-equation and system approaches. Even though they generate similar calibrations of the π_f, they generate very different predictions for the Vanek equation.

5.5. *Puzzle #2: Why Is The Formula for π^V_f So Odd?*

At first glance, the equation 6 expression for π^V_{fi} makes little sense. However, it turns out to be intimately related to results from the Development Accounting literature (Caselli, 2005, Caselli and Coleman, 2006, Caselli and Feyrer, 2007). To see this, consider a single-good economy with an aggregate production function $Y_i = \left[(\pi^V_{Li} V_{Li})^\sigma + (\pi^V_{Hi} V_{Hi})^\sigma\right]^{1/\sigma}$ where L is labor, H is human capital and Y_i is both output and income (the aggregate price is normalized to unity). The marginal product of factor f is $MP_{fi} = (\pi^V_{fi})^\sigma\,(Y_i/V_{fi})^{1-\sigma}$. Equating factor prices with marginal products, dividing by the corresponding equation for the United States, and substituting in PFPE yields

$$\pi_{fi}^{V} = \frac{s_i / V_{fi}}{s_{US} / V_{f,US}} \ .$$

Thus, the equation (6) expression for π_{fi}^{V} is the open-economy extension of a standard development accounting expression for factor augmenting productivity. π_{fi}^{V} thus links our work with this literature in which relative productivity levels are calculated by comparing relative outputs Y_i / Y_{us} with relative inputs $V_{fi} / V_{f,US}$.[18]

5.6. *Puzzle #3: Does Trade Matter?*

If we set the factor content data F_{fi}^{*} to 0, we obtain a seeming puzzle that $\widehat{\pi}_f$ dos not change much. However, the puzzle can now be easily resolved. Missing trade implies that small changes in π_f lead to large changes in $F_{fi}^{*}(\pi_f)$. Since $F_{fi}^{*}(\pi_f)$ is an invertible function, it follows that large changes in F_{fi}^{*} (setting it to 0) lead to small changes in π_f.

It is not surprising that trade does not matter very much in inferring productivity levels. In quantitative trade models [e.g. Arkolakis, Costinot and Rodriguez-Clare (2012)] the formula comparing welfare in one country to welfare in the United States is similar to π_L^V and we know that the welfare costs of setting trade to zero (autarky) from current levels of trade are small (e.g. Waugh (2010)).

That said, trade does play some role. From equation (6), π_{fi}^{V} is greater than its closed-economy version if $F_{fi}^{*} / s_i > F_{f,US}^{*} / s_{US}$. That is, if after controlling for size, country i export more services of factor f than does the United States, then country i is revealed by trade to be relatively productive in i. Observed trade matters for inferring productivity differences, but not by very much.

5.7. *Puzzle #4: The Gabaix Critique*

Gabaix (1997b) has nevertheless become viewed as one of the most withering criticisms of Trefler (1993). By implication, it is a criticism of our approach. Rather than engaging in a he-said-

[18] Here are the details. Setting $w_{fi} = MP_{fi}$ and dividing through by $MP_{f,US}$ yields $w_{fi}/w_{f,US} = MP_{fi}/MP_{f,US} = (\pi_{fi}^{V}/\pi_{f,US}^{V})^{\sigma} \left[(s_i/V_{fi})/(s_{US}/V_{f,US})\right]^{1-\sigma}$ where we have substituted in $Y_i = s_i Y_w$. Substituting in PFPE yields $\pi_{fi}^{V}/\pi_{f,US}^{V} = (\pi_{fi}^{V}/\pi_{f,US}^{V})^{\sigma} \left[(s_i/V_{fi})/(s_{US}/V_{f,US})\right]^{1-\sigma}$, from which the equation follows. Under the assumption of PFPE and no capital inputs, this is exactly the same prediction as in Caselli and Coleman (2006). With PFPE, there are no differences in factor wage shares across countries such that the expression in their footnote 7 collapses to the expression above when abstracting from capital.

she-said discussion, we present what we think are the three enduring criticisms in the folklore. The first and most important is that when the π_f^W are put into the Vanek equation, the fit is horrible. One might conclude from this that the model is wrong. However, we have provided two complementary explanations of this: (*a*) Missing trade implies that $F_{fi}^*(\pi_f)$ is sensitive and (*b*) single-equation approaches are lexicographic and hence inadequate. Second, Gabaix recalculates π_f^V by replacing the data F_{fi}^* with the data F_{fi}^* in the Vanek equation and finds that the new π_f is very similar to π_f^V. The implication is that trade plays no role and hence that the model is fundamentally wrong. As noted above, trade plays a role but as in many trade models, it is a small one. Third, the formula for π_f^V is 'rigged' to yield good results. Again as noted above, the formula is exactly what one finds in other literatures such as the Development Accounting literature — there is nothing unusual about it.

6. Empirical Specification with the Failure of PFPE

We have seen that there are empirical problems with the wage equation. Further, Davis and Weinstein (2001) argue that the failure of factor price equalization is an essential feature of the data. To investigate, we develop the simplest possible model that features a failure of PFPE while at the same time nesting our empirically successful Vanek and technology equations. Recall that the data underlying the technology equation is centred on the ratio $d_{fgi}^{-1}/d_{fg,US}^{-1}$, which means that we will need a diversified equilibrium in which good g is produced by multiple countries.

We make just one modification to the previous model. We now assume that the cost function for primary factors is

$$c_{gi}(\mathbf{w_i}) = \left[\sum_f \frac{\alpha_{fg}}{\delta_{gi}} \left(\frac{w_{fi}}{\pi_{fi}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} . \tag{8}$$

where $\alpha_{fg} > 0$ control for the factor intensity of the good and δ_{gi} is a Ricardian parameter that plays the central role of causing PFPE to fail.[19] It is easy to find a set of δ_{gi} that support a diversified equilibrium without PFPE. However, because the requirement of a diversified equilibrium imposes restrictions on the δ_{gi}, we will need to be careful in calibrating the δ_{gi} to a diversified equilibrium.

[19]With more goods than factors, all countries producing all goods, and all countries charging the same price for a given good, the zero profit conditions equation [2] show that common world goods prices pin down common productivity adjusted factor prices.

There are several minor points about diversification. First, the δ_{gi} can be interpreted as quality, in which case our diversification has the flavour of Schott (2003, 2004). Schott provides abundant evidence of diversification in his analysis of 'product overlap' at the 10-digit HS level. See also Sutton and Trefler (forthcoming). Second, we treat observed diversifcation as trade in varieties rather than as a function of aggregation bias. We do this both because of Schott's evidence and because this leads to a tight connection between our theory and empirics.[20] Third, country-level productivity can be loaded onto either the π_{fi} or the δ_{gi} so a normalization is needed. For comparability with the discussion of the PFPE case in the previous sub-section we normalize the δ_{gi} using $\delta_{g,US} = 1$ for all g and $\Sigma_g \theta_{Lgi}\delta_{gi} = 1\ \forall i$ where, as before, θ_{Lgi} is the share of country i's labor endowment (V_{Li}) employed in industry g.

If varieties of good g are produced both by country i and by the United States then Shephard's lemma implies

$$d_{fgi} = (\beta_{fi}\delta_{gi})^{-1} d_{fg,US} \tag{9}$$

where

$$\beta_{fi} \equiv \left(\frac{w_{fi}/\pi_{fi}}{w_{f,US}/\pi_{f,US}}\right)^{\sigma} \frac{\pi_{fi}}{\pi_{f,US}}\ . \tag{10}$$

See lemma 3 in Appendix A for a proof. It follows that $\overline{d}_{fi}^{-1} \equiv \sum_g \theta_{fgi} d_{fg,US}/d_{fgi}$ of equation (T) becomes

$$\overline{d}_{fi}^{-1} = \beta_{fi}\delta'_{fi} \tag{T'}$$

where $\delta'_{fi} \equiv \sum_{g=1}^{G} \theta_{fgi}\delta_{gi}$ aggregates sectoral-level Ricardian productivity differences to the factor level. Under our equation (8) assumption, the Vanek equation becomes

$$F^*_{fi} = \beta_{fi}\,\delta'_{fi}\,V_{fi} - s_i \sum_{j=1}^{N} \beta_{fj}\,\delta'_{fj}\,V_{fj}. \tag{V'}$$

Substituting factor demands (equation 9) into the factor-market clearing condition (equation 4) and solving for wages yields

$$\frac{w_{fi}/\pi_{fi}}{w_{f,US}/\pi_{f,US}} = \left[\frac{\pi_{f,US}V_{f,US}}{\pi_{fi}V_{fi}}\right]^{1/\sigma} \left(\sum_{g=1}^{G} \frac{d_{fg,US}Q_{gi}}{\delta_{gi}V_{fUS}}\right)^{1/\sigma}\ .$$

See Appendix A, lemma 4 for a proof. The first term in square brackets shows that productivity-adjusted factor prices are decreasing in productivity-adjusted factor supplies, *ceteris paribus*. The second term shows that the price of factor f is bid up if output Q_{gi} is large in sectors with high

[20] Davis and Weinstein (2001) and Feenstra and Hanson (2000) find it more useful in their contexts to treat observed diversification as a consequence of aggregation in an equilibrium with specialization.

per-unit demands for factor f, *ceteris paribus*. That is, these demands are high when the sector is intensive in factor f ($d_{fg,US}$ large) or unproductive (δ_{gi} small).[21] Rearranging this equation yields our third and final estimating equation:

$$\sum_{g=1}^{G} \frac{d_{fg,US} Q_{gi}}{V_{fi} \delta_{gi}} = \beta_{fi} \; . \tag{W'}$$

Finally, we calibrate the δ_{gi} using

$$\widehat{\delta}_{gi} = \frac{(d_{Lg,US}/d_{Lgi})^{1/2} (d_{Hg,US}/d_{Hgi})^{1/2}}{\sum_{g'=1}^{G} \theta_{Lg'i} (d_{Lg',US}/d_{Lg'i})^{1/2} (d_{Hg',US}/d_{Hg'i})^{1/2}} \; .$$

Using equation (9) to substitute out the $d_{fg,US}/d_{fgi}$ demonstrates that $\widehat{\delta}_{gi} = \delta_{gi}$.[22] Intuitively, these Ricardian differences are average differences in input requirements once the aggregate factor-specific influences of factor augmenting productivity and factor prices have been purged out.

6.1. Identification

In equations (T'), (V') and (W') the only unknown parameters are the β_{fi}. Further, the only place where w_{fi}, π_{fi} and σ appear are in the β_{fi}. Hence, given data on factor prices and estimated β_{fi},

$$\widehat{\beta}_{fi} = \left(\frac{w_{fi}/\pi_{fi}}{w_{f,US}/\pi_{f,US}} \right)^{\sigma} \frac{\pi_{fi}}{\pi_{f,US}} \; , \tag{11}$$

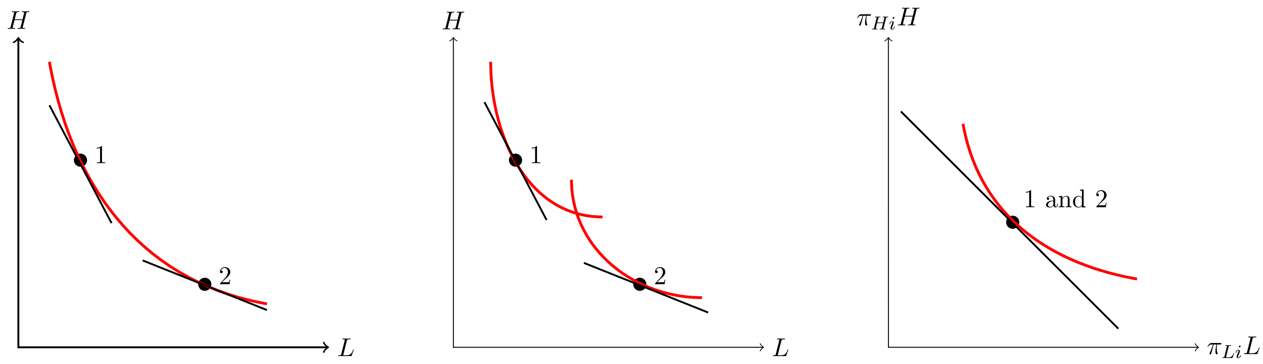
we can only estimate combinations of σ and the π_{fi}.

This lack of identification is not particular to our model. As Diamond et al. (1978) showed, it appears whenever cost and demand functions are estimated. To understand why, suppose that we observe data on factor prices and the amounts of L and H per unit of output used in two different countries i.e., suppose we observe (w_{Li}, w_{Hi}) and (d_{Lgi}, d_{Hgi}) for $i = 1,2$. These appear as the two slopes and two points in panel (*a*) of figure 4 and again in panel (*b*). Now consider the problem of estimating cost or demand functions that are consistent with these data. One approach is to make the identifying assumption that technologies are internationally identical and then to fit the data by adjusting the the curvature of the isoquant. See panel (*a*). In our CES context this means adjusting σ. Another approach is to assume that there are international technology differences so that isoquants differ across countries. See panel (*b*). In between there are countless other

[21]Simple intuition comes from the case of PFPE. All differences in relative factor supply (the term in square brackets) are exactly offset by differences in relative demand.

[22]As an aside, this normalization implies that $\delta'_{Li} = 1$. This in turn means that for $f = L$, equation (V) is the same as equation (V') but with β_{fi} taking the place of π_{fi}. Likewise for equations (T) and (T').

Figure 4: Identification



(*a*): Identical technologies
Choose curvature

(*b*): Identical curvature
Choose tech. diff.

(*c*): Technology differences
with PFPE

possibilities involving mixtures of curvature and international technology differences. Our model with PFPE is a very special case of this. All international technology differences are assumed to be factor augmenting (π_{fi}) and, when factors are put in productivity-adjusted units, factor prices are equalized. This is illustrated in panel (*c*) where the axes are productivity-adjusted factor inputs so that international differences in technology and factor prices disappear. Thus, our inability to separately estimate σ and the π_{fi} is a generic problem when estimating cost and production functions with international data.[23] Theorem 3 states this result formally.

Theorem 3 *Given data on unit input requirements* (d_{fgi})*, the factor content of trade* $(F_{fi,US})$*, production* (Q_{gi})*, endowments* (V_{fi})*, Ricardian productivity differences* $(\delta_{gi})^{1-\sigma}$*, country expenditure shares* (s_i)*, and wages* (w_{fi})*,*

1. *The vector* $\beta_f \equiv (\beta_{f1} \cdots \beta_{fN})$ *is overidentified, and*
2. *We cannot separately identify* σ *and the vector* π_f*.*

Proof See Appendix D. ■

Given these identification issues our strategy is as follows. We first show that the model without PFPE fits very well. We then choose a value of σ from the literature and use the $\widehat{\beta}_{fi}$ to back out the π_{fi}. Finally, we examine whether the π_{fi} are consistent with PFPE and assess the relative importance of (1) departures from PFPE and (2) international productivity differences for understanding the Vanek, wage, and technology predictions.

[23]In the special case where PFPE is assumed, all countries have the same factor prices so substitution effects disappear i.e., curvature plays no role. This is why in the case of PFPE we can recover the π_{fi} from the $\widehat{\beta}_{fi}$ even if σ is not known. That is, the first right-hand term in equation 11 vanishes. This explains why we did not make any reference to σ until now.

6.2. *Results with Failure of PFPE*

We now use our system of equations approach to jointly calibrate equations V', W', and T'. Figure 5 presents the results. It is identical to figure 2 except that $\widehat{\pi}_f$ is replaced by $\widehat{\beta}_f$. The overall impression is that the results are almost identical to what we saw in figure 2. That is, adding a failure of FPE makes no difference to the results. This is to be expected for the Vanek and technology equations since, as we noted, for labour there is no difference between the two specifications because $\delta'_{fi} = 1$. The R^2s for the Vanek, wage, and technology equations are 0.943, 0.943, and 0.949 for labor and 0.888, 0.888, and 0.901 for human capital.

The more interesting result appears when we difference the results across labor and human capital. This is illustrated in figure 6. It displays a huge improvement in the wage equation relative to what we saw with PFPE (figure 3).[24]

6.3. *Puzzle #5*

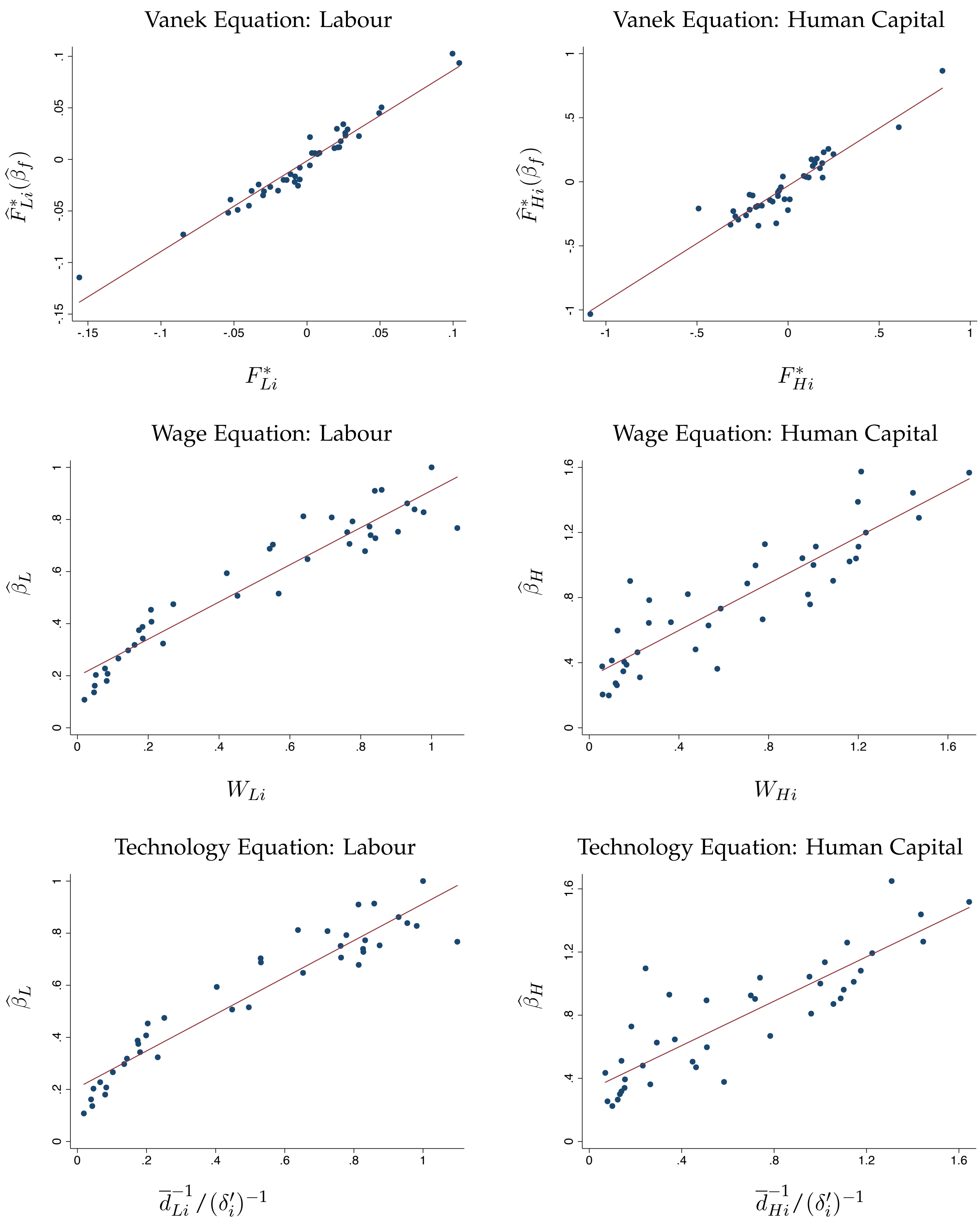
Why is this important? For one, Davis and Weinstein (2001) use their estimates to reject Trefler's (1993, 1995) model and, by implication, our model. They reject it in favour of a model featuring the failure of PFPE. Yet nowhere do they explicitly make any identifying assumptions outside of the hypothesis that there should be a reduced form relationship between aggregate relative endowments and relative input techniques. One of the primary goals of this section is to build on their work to offer a proper micro-founded description of the determinants of relative input techniques.

6.4. *Non-Traded Services*

It is remarkable that we are above to obtain such success with the model featuring a failure of PFPE despite the fact that we have abstracted from trade costs. The natural question is what would occur if we included trade costs as well. Because the existence of the Vanek equation (our trade equation) depends crucially on there being no trade costs, we step away from the model and follow Davis and Weinstein (2001) and examine the importance of non-traded services. For parsimony, we focus on one sector: government services. We focus on this sector for three

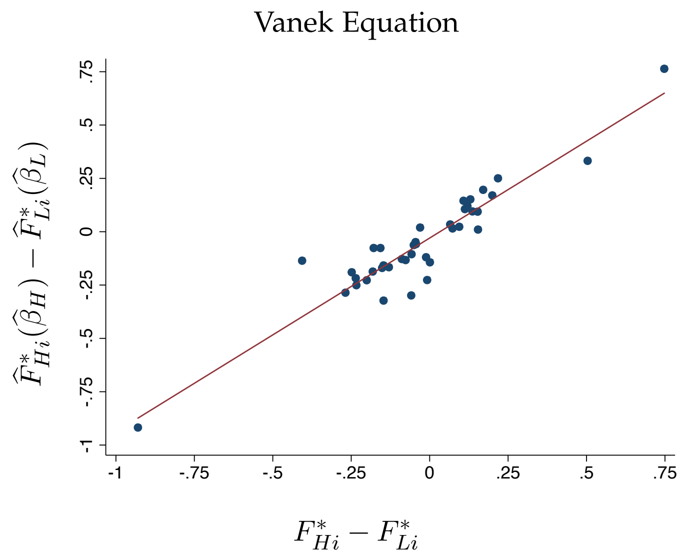
[24]Changing the objective function to include correlation coefficients instead of R^2 statistics delivers very similar results.

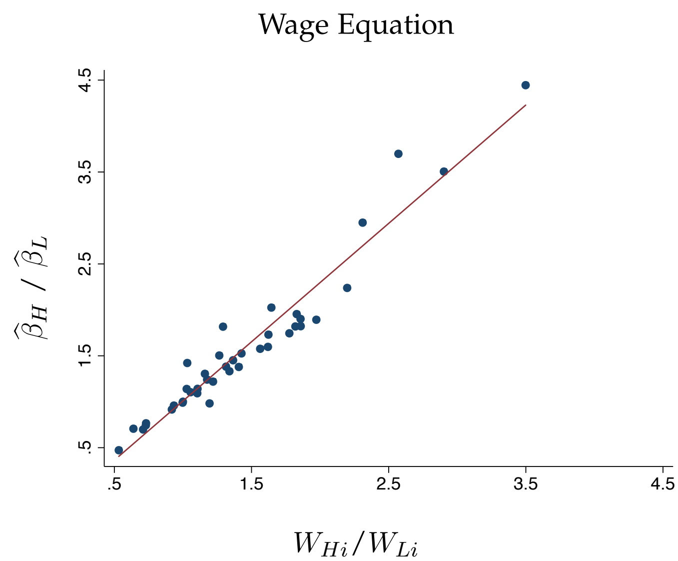
Figure 5: Failure of PFPE: Performance of the Vanek, Wage and Technology Equations

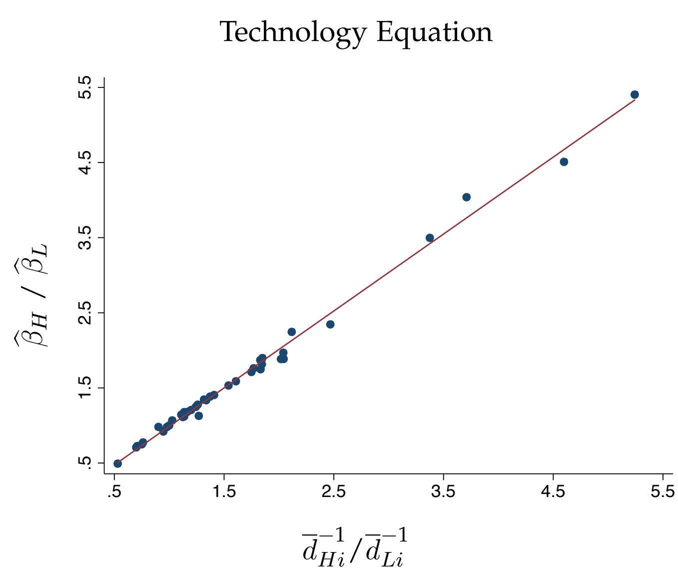


Notes: The top row plots the predicted productivity-adjusted factor content of trade $\widehat{F}^*_{fi}(\pi_f)$ from equation (7) against the actual productivity-adjusted factor content of trade F^*_{fi}. The middle row plots $W_{fi} \equiv \left(\frac{V_{fi}}{V_{f,US}}\right)\sum_{g=1}^{G}(d_{fg,US}/\delta_{gi})Q_{gi}$ against β_{fi}. The bottom row plots inverse (average) unit input requirements $(d^{-1}_{fi}/d^{-1}_{f,US})$ against β_{fi}. The left column is for labour and the right column is for human capital. Each row uses a the same $\hat{\pi}_f$ obtained as discussed in the text.

Figure 6: Failure of PFPE: Differencing Across Factors







reasons. First, the definition of government services varies dramatically across borders leading to the possibility of severe aggregation bias.[25] Second, the average shares of aggregate output value and value added are 13.8% and 19.2%, respectively, the largest of any sector in our sample. Consequently, its treatment can have a strong effect on aggregate outcomes. Third, unlike FIRE or transportation, there is little controversy that government services are generally non-traded.

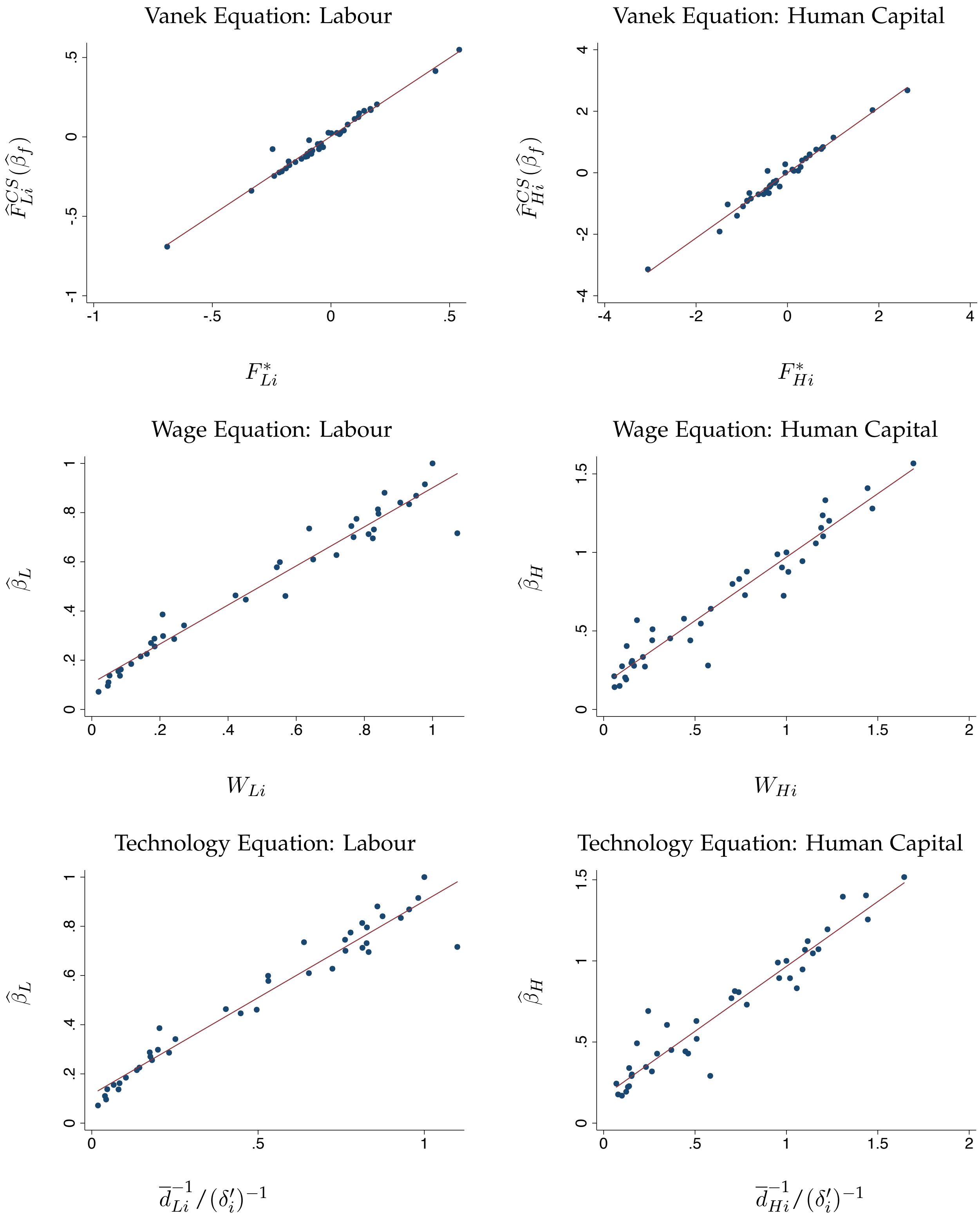
Following Trefler and Zhu (2010), we consider the following experiment. Suppose that government services were freely traded and that preferences were identical and homothetic. Then each country's consumption of a source country's government services would be equal to their aggregate income share times the source country's production of that good ($c_{gij} = s_i c_{g,world,j}$). Net trade would then be actual production minus this hypothetical set of consumption values (adjusting for intermediate input use). Appendix E discusses this in detail.

With these trade flows that are hypothetical *for government services alone*, we recalculate the factor content of trade and redo the calibration exercise. Figures 7 and 8 present our results visually. Figures 7 and 8 offer noticeable but modest improvements over figures 5 and 6.[26] Row 3 of table 1 presents test statistics. However, these improvements are second order relative to the strong performance of our model featuring a failure of PFPE.

7. Productivity

We have thus far calculated technology terms β_{fi} that fit data on the factor content of trade, relative factor inputs, and relative factor market clearing when PFPE fails. These appear in columns (1) and (5) of table 3. However, they are reduced form combinations of equilibrium factor prices, factor augmenting productivity, and the elasticity of substitution in production and do not identify one mechanism relative to another. We now exploit our data on factor prices and calibrate an elasticity of substitution to extract values of factor augmenting productivity. We emphasize that our results up to this point for the case in which PFPE fails do not depend on the precise values of productivity that we calculate here and that the structural combination β_{fi}

Figure 7: Adjusting for Non-Traded Government Services



Notes: The top row plots the predicted productivity-adjusted factor content of trade adjusted for consumption similarity in government services, $\widehat{F}^{CS}_{fi}(\pi_f)$ from equation (7) against the actual productivity-adjusted factor content of trade F^{CS}_{fi}. The middle row plots $W_{fi} \equiv \left(\frac{V_{fi}}{V_{f,US}}\right) \sum_{g=1}^{G}(d_{fg,US}/\delta_{gi})Q_{gi}$ against β_{fi}. The bottom row plots inverse (average) unit input requirements $(d_{fi}/d_{f,US})$ against β_{fi}. The left column is for labour and the right column is for human capital. Each row uses a the same $\hat{\pi}_f$ obtained as discussed in the text.

Figure 8: Failure of PFPE and Consumption Similarity: Differencing Across Factors

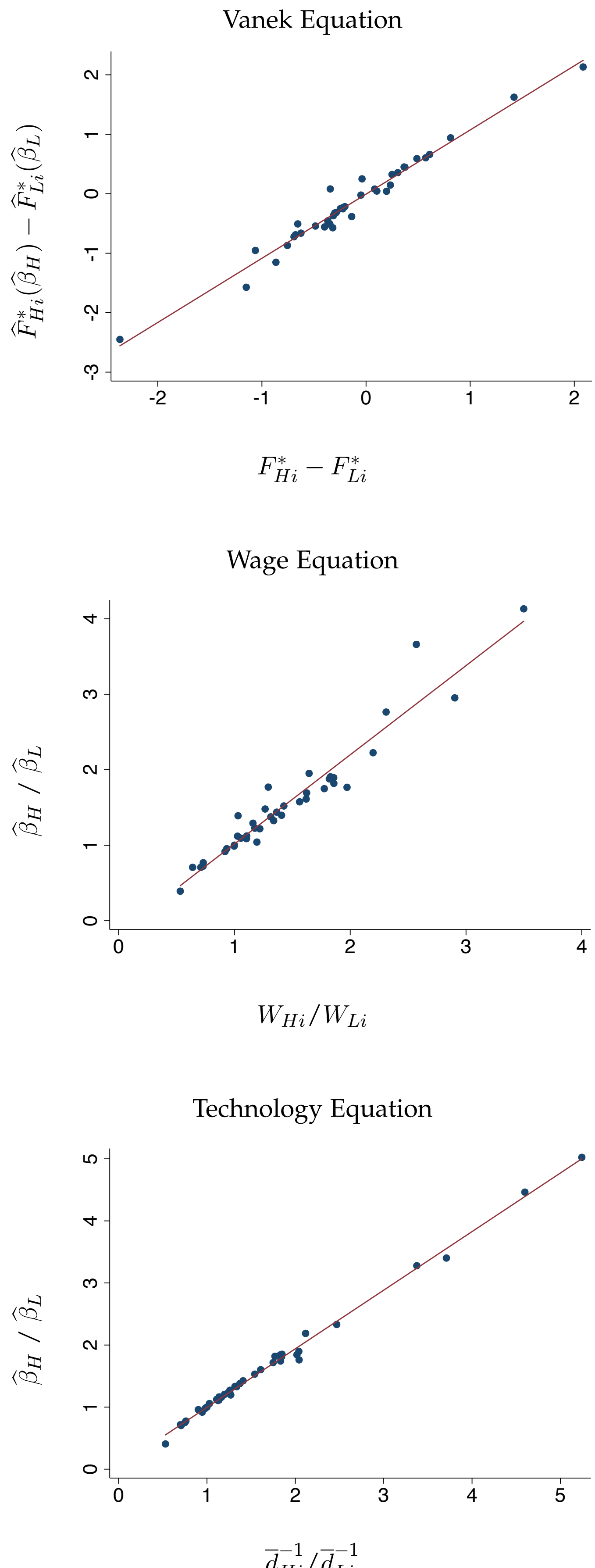


Table 3: Calibrated Values of β_{fi} and π_{fi}

	Labor				Human Capital			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	$\hat{\beta}_{Li}$	w_{Li}	$\hat{\pi}_{Li}$	$\left(\frac{w_{Li}}{\hat{\pi}_{Li}}\right)^{\sigma}$	$\hat{\beta}_{Hi}$	w_{Hi}	$\hat{\pi}_{H,i}$	$\left(\frac{w_{Hi}}{\hat{\pi}_{Hi}}\right)^{\sigma}$
Argentina	0.453	0.289	0.557	0.814	0.347	0.390	0.329	1.055
Australia	0.808	0.997	0.734	1.101	1.113	1.023	1.157	0.962
Austria	0.753	0.895	0.696	1.082	1.040	1.089	1.018	1.021
Belgium	0.862	1.160	0.752	1.146	1.567	1.418	1.641	0.955
Brazil	0.323	0.197	0.406	0.797	0.310	0.405	0.274	1.131
Canada	0.812	0.752	0.841	0.966	0.887	0.835	0.911	0.973
Chile	0.407	0.281	0.482	0.844	0.464	0.400	0.496	0.935
China	0.162	0.047	0.283	0.570	0.597	0.085	1.456	0.410
Colombia	0.207	0.180	0.221	0.938	0.405	0.364	0.425	0.953
Denmark	0.828	0.975	0.768	1.078	0.819	1.033	0.736	1.112
Finland	0.751	1.079	0.636	1.180	1.199	1.333	1.143	1.050
France	0.792	1.305	0.630	1.257	1.443	1.765	1.316	1.097
Germany	0.740	1.310	0.569	1.299	1.290	1.571	1.179	1.094
Great Britain	0.706	0.994	0.604	1.170	1.113	1.201	1.074	1.036
Greece	0.515	0.326	0.636	0.810	0.732	0.402	0.964	0.760
Hong Kong	0.767	0.296	1.187	0.646	0.362	0.331	0.378	0.959
Hungary	0.375	0.119	0.634	0.591	0.261	0.121	0.372	0.702
Indonesia	0.203	0.089	0.296	0.685	0.902	0.137	2.136	0.422
Ireland	0.914	0.816	0.962	0.950	1.042	0.869	1.132	0.920
Italy	0.910	0.894	0.917	0.992	1.388	0.929	1.668	0.832
Japan	0.678	1.114	0.540	1.255	0.903	1.380	0.744	1.214
Korea	0.593	0.499	0.643	0.923	1.128	0.608	1.497	0.753
Malaysia	0.474	0.215	0.681	0.696	0.821	0.284	1.335	0.615
Mexico	0.387	0.092	0.748	0.518	0.274	0.126	0.390	0.701
Netherlands	0.839	0.866	0.826	1.015	1.022	0.925	1.069	0.956
New Zealand	0.647	0.488	0.737	0.878	0.481	0.522	0.464	1.038
Peru	0.180	0.117	0.219	0.821	0.199	0.159	0.220	0.904
Phillipines	0.136	0.096	0.159	0.854	0.204	0.161	0.228	0.896
Poland	0.297	0.123	0.446	0.667	0.387	0.142	0.614	0.631
Portugal	0.507	0.296	0.648	0.782	0.628	0.428	0.750	0.838
Singapore	0.773	0.438	1.002	0.771	0.758	0.749	0.763	0.994
Spain	0.688	0.629	0.716	0.960	0.998	0.789	1.111	0.898
Sri Lanka	0.108	0.063	0.137	0.784	0.377	0.083	0.754	0.500
Sweden	0.728	0.887	0.665	1.095	0.666	0.908	0.578	1.152
Taiwan	0.703	0.817	0.657	1.071	1.574	1.101	1.854	0.849
Thailand	0.228	0.102	0.328	0.693	0.413	0.190	0.590	0.701
Turkey	0.266	0.312	0.247	1.076	0.784	0.524	0.943	0.832
Uruguay	0.343	0.335	0.346	0.990	0.649	0.487	0.740	0.877
USA	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Venezuela	0.318	0.151	0.448	0.711	0.644	0.260	0.976	0.660
Rank Corr. w/$gdpc$	0.931	0.831	0.829	0.567	0.614	0.771	0.296	0.605

Columns (1) and (5) present β_{fi} from the estimation in row (5) of table 1. Columns (2) and (6) are data described in section 4. Columns (3),(4),(7), and (8) are calculated using the data in columns (1), (2),(5),(6), the definition of β_{fi} and σ=.31.

is what is identified.

The intuition for how we can extract these values comes from Caselli and Coleman (2006). For a given value of σ, we can use our wage data to predict how much input techniques will vary across countries. Any residual variation is then attributed to differences in factor augmenting productivity. For example, suppose that some countries use a more labor intensive technique than would be expected given data on the elasticity of substitution across factors and factor price data. If factors of production are substitutes ($\sigma > 1$), then factor augmenting productivity must be higher for labor in these countries. If factors of production are complements ($\sigma < 1$), then factor augmenting productivity must be lower.

We first require a value of σ. While there are many values for the elasticity of substitution between skilled and unskilled labor, we know of none between human capital and labor.[27] We start by noting that our measure of human capital is high school equivalent workers as measured by total payroll divided by the wage of a high school graduate. We then represent skilled and unskilled workers in country i as V_{Si} and V_{Ui}, respectively, and assume that they are defined to be mutually exclusive with skilled workers possessing a high school education or above and unskilled workers with less.[28] The ratio of human capital to labor can then be written in terms of the ratio of skilled to unskilled labor and the relative wage of the two, w_{si}/w_{ui}: $\frac{V_{Hi}}{V_{Li}} = \frac{V_{Si}+\frac{w_{ui}}{w_{si}}V_{Ui}}{V_{Si}+V_{Ui}}$. Taking a total (log) derivative of this in terms of w_s/w_u and using the definition of the elasticity of substitution between skilled and unskilled labor, we obtain an expression for $\frac{\partial \ln \frac{V_{Hi}}{V_{Li}}}{\partial \ln \frac{w_s}{w_u}}$ that is a function of the V_{Si}/V_{Ui}, w_{si}/w_{ui}, and the elasticity of substitution between skilled and unskilled labor. We use a value of 1.4 for the elasticity of substitution between skilled and unskilled labor from Katz and Murphy (1992). Because the vast majority of estimates of this elasticity come from US data, we also use values of V_{Si}/V_{Ui} and w_{si}/w_{ui} from the US and obtain an (absolute) value

[25]For example, government services in Canada include healthcare and utilities, whereas in the US, each is commonly provided by private agents.

[26]The R^2 statistics for the trade, wage, and technology equations are 0.974, 0.974, and 0.976, respectively, for labor and 0.967, 0.967, and 0.969 for human capital.

[27] Ciccone and Peri (2005) review estimates of the elasticity of substitution between skilled and unskilled labor. Antràs (2004) and Oberfield and Raval (2014) analyze the elasticity of substitution between capital and labor.

[28]While the literature of estimating the elasticity of substitution between skilled and unskilled labor assumes that they are imperfectly substitutable, our framework assumes that they are perfectly substitutable in efficiency adjusted terms. This is an important caveat to this section.

of $\sigma = .31$.[29]

Using $\sigma = .31$, our calibrated values β_{fi}, and data for w_{fi}, we calculate and present values of π_{fi} in columns (3) and (7) of Table 3 and present $(w_{fi}/\pi_{fi})^{\sigma}$ in columns (4) and (8). The bottom row presents the rank correlation of each of these with real GDP per capita (y_i) from the Penn World Tables in 1997. Simply by inspecting columns (4) and (8), we reject the null hypothesis of productivity adjusted factor price equalization.

More interestingly, using our definition of β_{fi}, we can use this data gauge the importance of departures from PFPE, $(w_{fi}/\pi_{fi})^{\sigma}$, relative to differences in factor augmenting productivity π_{fi} in determining techniques. Examining column (1) of table 3, β_{fi} ranges between its minimum and maximum by a factor of 9.3. $(w_{fi}/\pi_{fi})^{\sigma}$ varies by a factor of 2.5 and π_{Li} varies by a factor of 8.7. For human capital, these factors are 7.9, 2.9, and 9.7. While departures from PFPE seem to play some role, it is minor with the majority of the variance accounted for by factor augmenting productivity.

8. Conclusion

Judging by current research in international economics, one can only conclude that models of factor abundance are wrong, useless, have had their intellectual content exhausted, or some combination of the three. There are three commonly cited reasons for this. First, Ricardian differences are important. Second, unit factor requirements differ across countries because of either international technology differences or the failure of productivity-adjusted factor price equalization. Third, budget shares are not the same across countries perhaps due to transport costs or non-homotheticities.

To assess these possibilities, this paper lays out the entire empirical content of a general equilibrium model of relative factor abundance. Using cross-equation restrictions, we start by calibrating international differences in factor augmenting productivity for a model in which PFPE holds such that all differences in relative wages perfectly reveal differences in factor augmenting productivity. The model provides a superb fit for data describing approximately 90% of the variation in equations that describe the factor content of trade, relative differences in input

[29]We use $V_{S,US}/V_{U,US}$=2.64 and and $w_{s,US}/w_{u,US}$=1.77. The first is adjusted for efficiency units *within* each group as in Caselli and Coleman (2006) (e.g. the stock of college educated is transformed into the number of equivalent high school educated). We have also examined the case where our σ is based on an elasticity of substitution between skilled and unskilled labor of 1.7 as in Krussell et al. (2000). Results are very similar.

techniques, and relative wages when examining the model factor-by-factor. However, when we difference our results across factors to assess issues of *relative* factor abundance, we find that relative differencs factor augmenting productivity across human capital and labor only weakly correlate with the observed relative human capital-labor wage gap.

We then examine our model in a setting in which PFPE fails due to the inclusion of Ricardian terms. Following the insights of Diamond, McFadden, and Rodríguez (1978), we show how one can only identify reduced form technology terms that embody the joint effect of differences in factor augmenting productivity, factor prices, and the elasticity of substitution when PFPE fails. While the model fit is better than when we assume that PFPE holds, the improvement is minor. We also find that considering non-traded government services improves the fit of the model but only by a small margin. We also show how to extract values of factor augmenting productivity from these reduced form technology parameters.

Appendix A. Mathematical Proofs

Lemma 1 *$q_{gi} = p_{gi}^{-\rho_g}\kappa_g$ for some $k_g > 0$ and all i.*

Proof From equation (1), $\phi_{hj}(\nu,\mathbf{w_j},\mathbf{p}) = [c_{hj}(w_j)]^{\gamma_{h0}} \prod_{g'=1}^{G} [P_{hg'}(\nu)]^{\gamma_{hg'}}$ where $P_{hg}(\nu) = \left(\sum_{i=1}^{N} \int_{\omega\in\Omega_{gi}} \alpha_{hg}(\nu,\omega)\, p_{gi}(\omega)^{1-\rho_g}\, d\omega \right)^{\frac{1}{1-\rho_g}}$. Note that $\partial P_{hg'}(\omega)/\partial p_{gi}(\omega) = 0$ for $g' \neq g$ and $\partial P_{hg}(\nu)/\partial p_{gi}(\omega) = P_{hg}(\nu)^{\rho_g}(1-\rho_g)\alpha_{hg}(\nu,\omega)p_{gi}(\omega)^{-\rho_g}$. Hence $\partial\phi_{hj}(\nu,\mathbf{w_j},\mathbf{p})/\partial p_{gi}(\omega) = c_{hj}^{\gamma_{g0}}\left[\Pi_{g'\neq g}P_{hg'}(\nu)^{\gamma_{g'h}}\right]\left[P_{hg}(\nu)^{\gamma_{hg}-1+\rho_g}(1-\rho_g)\alpha_{hg}(\nu,\omega)p_{gi}(\omega)^{-\rho_g}\right]$. Demand for variety $\omega \in \Omega_{gi}$ is the sum of demands for final goods and intermediate inputs: $q_{gi}(\omega) = p_{gi}(\omega)^{-\rho_g}P_g^{\rho_g-1}\eta_g\Sigma_{j=1}^{N}Y_j + \Sigma_{h=1}^{G}\Sigma_{j=1}^{N}\int_{\nu\in\Omega_{hj}} b_{ij}(g,h)\left[q_{hj}(\nu)+\overline{\phi}_h\right]d\nu$ where $b_{ij}(g,h) = \partial\phi_{hj}(\nu,\mathbf{w_j},\mathbf{p})/\partial p_{gi}(\omega)$ is variety ν's demand for ω. Hence $b_{ij}(g,h) = \phi_{hj}(\nu,\mathbf{w_j},\mathbf{p})\; P_{hg}(\nu)^{\rho_g-1}(1-\rho_g)\alpha_{hg}(\nu,\omega)p_{gi}(\omega)^{-\rho_g}$. Thus, the term $p_{gi}(\omega\,)^{-\rho_g}$ factors out of each term in $q_{gi}(\omega)$ and the remaining terms are independent of i. ■

Theorem 1 *$F_{fi} = V_i - s_i\Sigma_j V_{fj}$.*

Proof Pre-multiplying equation (5) by $\mathbf{A_f}$ yields $\mathbf{A_f T} = \mathbf{A_f}(\mathbf{I_{NG}}-\mathbf{B})^{-1}\mathbf{Q} - \mathbf{A_f C} = \mathbf{D_f Q} - \mathbf{A_f C} = \left[\begin{array}{ccc} V_{f1} & \cdots & V_{fN}\end{array}\right] - \mathbf{A_f C}$. Consider column i of this equation, namely,

$$\mathbf{A_f T_i} = V_{fi} - \mathbf{A_f C_i} \tag{12}$$

where $\mathbf{T_i}$ and $\mathbf{C_i}$ are the ith columns of T and C, respectively. Then

$$\mathbf{A_f}\Sigma_j\mathbf{T_j} = \Sigma_j V_{fj} - \mathbf{A_f}\Sigma_j\mathbf{C_j}. \tag{13}$$

Consider each of the three terms in this equation. $V_{fw} \equiv \Sigma_j V_{fj}$ is the world endowment of f. Recall that $\mathbf{T_j}$ is composed of blocks of $G\times 1$ matrices. Let $\mathbf{T_{ij}}$ be the ith block of $\mathbf{T_j}$. Then by inspection of the definition of $\mathbf{T}$, $\Sigma_j\mathbf{T_{ij}} = \mathbf{X_i} - \Sigma_j\mathbf{M_{ji}} = \mathbf{0_G}$ where $\mathbf{0_G}$ is the $G\times 1$ vector of zeros. Hence $\Sigma_j\mathbf{T_j} = \mathbf{0_{NG}}$ where $\mathbf{0_{NG}}$ is the $NG\times 1$ vector of zeros. Likewise, $\mathbf{C_j}$ is composed of blocks of $G\times 1$ matrices and $\mathbf{C_{ji}}$ is the ith block of $\mathbf{C_j}$. $\Sigma_j\mathbf{C_{ji}}$ is world consumption of good g produced in country i. Define $\mathbf{C_{wi}} = \Sigma_j\mathbf{C_{ji}}$ and stack the $\mathbf{C_{wi}}$ into an $NG\times 1$ vector denoted by $\mathbf{C_w}$. Thus, equation (13) can be written as $0 = V_{fw} - \mathbf{A_f C_w}$ or $0 = s_iV_{fw} - \mathbf{A_f}(s_i\mathbf{C_w})$. Subtracting this from equation (12) yields $F_{fi} = V_{fi} - s_iV_{fw} - \mathbf{A_f}\left(\mathbf{C_i} - s_i\mathbf{C_w}\right)$. But under our assumptions of homothetic demands and costless trade, $C_{gij} = s_iC_{gwj}$ or, in matrix notation, $\mathbf{C_i} = s_i\mathbf{C_w}$. Hence $F_{fi} = V_{fi} - s_iV_{fw}$. ■

Lemma 2 *(1) $\pi_{fi}d_{fgi}(\mathbf{w_i},\mathbf{p}) = d^*_{fg}(w^*_f,\mathbf{p})$ and (2) $F^*_{fi} = \pi_{fi}V_{fi} - s_i\Sigma_j\pi_{fj}V_{fj}$.*

Proof Part 1: Consider equation (1) and define $\phi^*_g(\mathbf{w_i}^*,\mathbf{p}) = \left[c^*_g(w_{1i}/\pi_{f1},\ldots,w_{Ki}/\pi_{Ki})\right]^{\gamma_{g0}}\prod_{h=1}^{G}[P_{gh}(\omega)]^{\gamma_{gh}}$. Then $\phi_{gi}(\mathbf{w_i},\mathbf{p}) = \phi^*_g(\mathbf{w_i}^*,\mathbf{p})$. By Shephard's lemma, $d_{fgi}(\mathbf{w_i},\mathbf{p}) = \partial\phi_{gi}(\mathbf{w_i},\mathbf{p})/\partial w_{fi}$ and $d^*_{fg}(\mathbf{w_i}^*,\mathbf{p}) = \partial\phi^*_g(\mathbf{w_i}^*,\mathbf{p})/\partial w_{fi}$. Hence, $d_{fgi}(\mathbf{w_i},\mathbf{p}) = \partial\phi^*_g((w_{1i}/\pi_{1i},\ldots,w_{Ki}/\pi_{Ki},\mathbf{p})/\partial(w_{fi}/\pi_{fi})\times\ \partial(w_{fi}/\pi_{fi})/\partial w_{fi} = d^*_{fg}(\mathbf{w_i}^*,\mathbf{p})\times\pi_{fi}$. That is, $\pi_{fi}d_{fgi}(\mathbf{w_i},\mathbf{p}) = d^*_{fg}(\mathbf{w_i}^*,\mathbf{p})$.

Part 2: The proof is alomst identical to the proof of theorem 1 except that D_f is replaced by D^*_f. This only affects the terms $D_fQ_i = V_{fi}$, which now become $D^*_fQ_i = \pi_{fi}V_{fi}$. Everything else goes through unchanged. ■

Lemma 3 *Equation (9)*

Table 4: Countries

Argentina	Australia	Austria	Belgium	Brazil	Canada	Chile	China
Colombia	Denmark	Finland	France	Germany	Great Britain	Greece	Hong Kong
Hungary	Indonesia	Ireland	Italy	Japan	Korea	Malaysia	Mexico
Netherlands	New Zealand	Peru	Phillipines	Poland	Portugal	Singapore	Spain
Sri Lanka	Sweden	Taiwan	Thailand	Turkey	United States	Uruguay	Venezuela

Table 5: Industries

ISIC Code	Industry Name	ISIC Code	Industry Name
110	Agriculture	371	Iron and Steel
200	Mining	372	Non-ferrous metals
311	Food	381	Fabricated metal products
313	Beverages	382	Machinery except electrical
321	Textiles	384	Transport Equipment
322	Wearing Apparel	390	Misc. Manufactures
323	Leather	400	Electricity, Gas, and Water
331	Wood Prod.	500	Construction
341	Paper and Paper Prod.	600	Wholesale Trade-Retail
351	Industrial Chemicals	700	Transport, Storage and Comm.
353	Petroleum	800	Financing, Insurance, Real Estate, etc.
361	Pottery	900	Government Servides

Proof $d_{fgi} = [\partial\phi_{gi}/\partial c_{gi}]\,[\partial c_{gi}/\partial w_{fi}] = \left[\gamma_{g0} c_{gi}^{-1} \phi_{gi}\right] \left[\alpha_{fg}\,(w_{fi})^{-\sigma}\,(\pi_{fi})^{\sigma-1}\,(\delta_{gi})^{\sigma-1}\,(c_{gi})^{\sigma}\right]$. Recall that in section 2.3 we established that $p_{gi} = p_g$. Hence from equation (2), $1 = p_{gi}/p_{g,US} = \phi_{gi}(\mathbf{w_i},\mathbf{p})/\phi_{g,US}(w_{US},\mathbf{p}) = \left(c_{gi}^{\gamma_{g0}} \Pi_{h=1}^{G} P_{gh}^{\gamma_{gh}}\right) / \left(c_{g,US}^{\gamma_{g0}} \Pi_{h=1}^{G} P_{gh}^{\gamma_{gh}}\right) = (c_{gi}/c_{g,US})^{\gamma_{g0}}$ or $c_{gi} = c_{g,US}$. Hence, in the ratio $d_{fgi} = d_{fg,US}(w_{fi}/w_{f,US})^{-\sigma}(\pi_{fi}/\pi_{f,US})^{\sigma-1}(\delta_{gi}/\delta_{g,US})^{\sigma-1}$. ■

Lemma 4 *Equation (W')*

Proof Plugging $Q_{gi} = n_{gi}(q_g + \overline{\phi}_g)$ into the factor-market clearing equation (4) yields $V_{fi} = \Sigma_g d_{fgi} Q_{gi}$ or $V_{fi} = \frac{d_{fgi}}{d_{fj,US}} d_{fg,US} Q_{gi}$. Plugging equation (9) into this equality yields $V_{fi} = (w_{fi}/w_{f,US})^{-\sigma}(\pi_{fi}/\pi_{f,US})^{\sigma-1}\Sigma_g d_{fg,US} Q_{gi}(\delta_{gi}/\delta_{g,US})^{\sigma-1}$. Using $\theta_{fg,US} = d_{fg,US} Q_{g,US}/V_{f,US}$ to substiute out $d_{fg,US}$ yields equation (W). ■

Appendix B. Data Appendix

Sample

Countries and industries used are given in tables 4 and 5.

Construction of World Input-Output Matrix B

We now describe in detail our construction of the world input-output matrix **B**. Recall that it is defined as

Table 6: Occupations

Occ. Code	Occ. Name	Occ. Code	Occ. Name
20	Butcher	31	Tanner
21	Packer	32	Leather goods maker
22	Dairy processor	33	Clicker cutter
23	Grain miller	34	Laster
24	Baker	34	Shoe sewer
25	Thread spinner	36	Sawmill sawyer
26	Loom fixer	37	Veneer cutter
27	Cloth weaver	38	Plywood press operator
28	Labourer	39	Furniture upholsterer
29	Garment cutter	40	Cabinetmaker
30	Sewing-machine operator		

$$\mathbf{B} \equiv \begin{bmatrix} \mathbf{B_{11}} & \mathbf{B_{12}} & \cdots & \mathbf{B_{1N}} \\ \mathbf{B_{21}} & \mathbf{B_{22}} & \cdots & \mathbf{B_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B_{N1}} & \mathbf{B_{N2}} & \cdots & \mathbf{B_{NN}} \end{bmatrix}.$$

Because each country faces the same set of intermediate input prices, each producing country uses the same techniques. Consequently, each row is identical such that $\mathbf{B}_{ij} = \mathbf{B}_{ij'}\ \forall i,j,j'$. We consequently start by using reported US data from GTAP to construct our measure of $\mathbf{B}_{ij}$. Unfortunately, the best reported data does not report use of intermediate inputs by source country (e.g. how $\mathbf{B_{ij}}$ varies from $\mathbf{B_{i'j}}$.[30]). To keep our data logically consistent with our theoretical framework, we assume that $\frac{b_{ij}(g,h)}{b_{i'j}(g,h)} = \frac{Q_{gi}}{Q_{gi'}}$ which is consistent with our model's result of consumption similarity.

Construction of Labor Variables

To obtain the wage rate of labor, we use wage rate of occupations 20-40 in the ILO October Inquiry to represent the wage rate of labor. Occupations are listed below

This allows us to pick our occupations that are generally considered unskilled and in highly traded industries. It also allows us to average out idiosyncratic country-occupation factors that may introduce measurement error (e.g. sewing machine operators in Italy). All wage data are in nominal monthly USD wages. Because our trade and production data are for 1997, we use wage data for this year. For observations for which data are not available in this year, we use the following algorithm: for the years 1995-2000 (inclusive) find years in which this country-occupation possesses data and so does the United States. Take the ratio of the two. Denote this value as $w_{oit}/w_{o,US,t}$ for occupation o for country i at time t. If there are O_i such observations for a given country, we then take the log of this $\ln(w_{oit}/w_{o,US,t})$ and take the mean of $mean(\ln(w_{oit}/w_{o,US,t})) = \frac{1}{O_i}\sum_{\ln(w_{oit}/w_{o,US,t})}$. We then exponentiate this value to obtain our final value $w_{L,i}/w_{L,US} = \exp\left[mean(\ln(w_{oit}/w_{o,US,t}))\right]$.

For some countries there are no such values over which to obtain this value. These countries are Chile, Colombia, Denmark, France, Greece, Indonesia, Ireland, the Netherlands, New Zealand, Spain, Turkey, and Taiwan. For these counties we impute values of $w_{L,i}/w_{L,US}$ as

[30]Data sets such as the World Input-Output Database use a proportionality assumption to allocate inputs across source countries.

follows. Take the OWW dataset for all countries but occupations 20-40 and the years 1995-2000 (inclusive). At the country-year level for each occupation o regress ln(w_{oit} on (log) real GDP per capital, (log) total employment, (log) labor share of income, (log) capital stock, manufacturing exports share of exports, manufacturing share of imports, (log) TFP, the square of each of these terms,region dummies (north America, south America, north Europe, south Europe, Africa, Asia, and Australia). Take the mean of the fitted values of these regressions for the years 1995-2000 across occupations, and exponentiate for the final imputed value. The median R^2 statistic for these regressions is .96 with a minimum of .92 and a maximum of .98. The median number of observations is 97.

To calculate human capital endowments and wages, start by letting V_{Lgi} and PL_{gi} represent physical labor employment and payroll of industry g in country i. Data are from the OECD STAN data base for OECD countries, the UNIDO data base for manufacturing in non-OECD countries, and from the ILO for non-manufacturing in non-OECD countries. The endowment of labor if $V_{Li} = \sum_g V_{Lgi}$.

The construction of human capital is as follows. The returns to education λ_i are from Banerjee and Duflo (BD, 2005) with the following exceptions: Belgium (missing from BD, use French value), Ireland (missing from BD, use English value), Mexico (.076 from P&P (2004); BD report .35), New Zealand (missing from BD, use value from P&P (2004)), Turkey (missing from BD, use value from P&P (2004)), and Taiwan (missing from BD, use value from P&P (2004)).

Assume that (1) $\sum_g PL_{gi} = \sum_e w_i(e) L_i e$, where e denotes years of schooling described below and $L_i(e)$ come from the Barro-Lee educational attainment data base; and (2) $w_i(e) = (1+\lambda_i)^e \, w_i(0)$. Combining these two assumptions yields $\sum_g PL_{gi} = \sum_e (1+\lambda_i)^e w(0) L_i(e)$. Thus

$$w_{gi}(0) = \frac{\sum_g PL_{gi}}{\sum_e (1+\lambda_i)^e L_{gi}(e)}.$$

Define

$$V_{Hgi} \equiv \frac{PL_{gi}}{w^c(12)} = \frac{PL_{gi}}{w^c(0)(1+\lambda_i)^{12}}.$$

Because our measure of human capital is the workforce measured in grade 12 equivalent workers, the wage of human capital in our empirical work is $w_i(12)$. Equation (Appendix B) shows how to calculate the country-industry-level measure of human capital V_{Hgi}. It converts the workforce in the unit of grade 12 equivalent (i.e. $\sum_e (1+\lambda_i)^{e-12} L_i(e)$) and then allocates the workforce total across industries using $PL_{gi} / \sum_{g'} PL_{g'i}$. Two more points can be made about this equation

1. V_{Hgi} is independent of $\mathbf{w_i}(e)$ including $w_i(0)$ and $w_i(12)$.
2. V_{Hgi} is independent of L_{gi}. Instead, constructing V_{Hgi} uses $L_i(e)$ from Barro-Lee. Thus to make V_{Hgi} and V_{Lgi} consistent, the endowment of labor $V_{Li} = \sum_g V_{Lgi}$ must be scaled so that it sums to the Barro and Lee workforce totals. On the other hand, we can still use the original payroll data PL_{gi} to allocate workers (in the unit of grade 12 equivalent) across industries.

Appendix C. Algorithm

Before we proceed to the main theorem and its proof, we discuss its preliminary set-up.

Appendix C.1. Setup

Consider three linear models indexed by k, each with N observations (countries) and N parameters:

$$Y_k = X_k\beta_k \quad k = 1,2,3$$

where $\beta_k = (X_k'X_k)^{-1}X_k'Y_k$ and the model fits perfectly. Let $e_k(b) = Y_k - X_kb$ be the residuals from the k^{th} equation when the vector of coefficients is b rather than β_k.

Define the unit simplex $\Delta \equiv \{a : a = (a_1,a_2,a_3),\ a_k \geq 0,\ \Sigma_k a_k = 1\}$ and stack the three equations:

$$Y(a) = \begin{bmatrix} \sqrt{a_1}Y_1 \\ \sqrt{a_2}Y_2 \\ \sqrt{a_3}Y_3 \end{bmatrix} \quad \text{and} \quad X(a) = \begin{bmatrix} \sqrt{a_1}X_1 \\ \sqrt{a_2}X_2 \\ \sqrt{a_3}X_3 \end{bmatrix}.$$

Let $e(b,a) = Y(a) - X(a)b$ be the residuals when we impose a single slope parameter b for all three equations. The minimizer of $e(b,a)'e(b,a)$, namely $b(a) = [X(a)'X(a)]^{-1} X(a)'Y(a)$, can be written as the matrix-weighted average of the b_k. To see this, note that

$$\begin{aligned} b(a) &= \left[X(a)'X(a)\right]^{-1} X(a)'Y(a) \\ &= \left[\Sigma_{k=1}^3 a_k X_k'X_k\right]^{-1} \left[\Sigma_{k=1}^3 a_k X_k'Y_k\right] \\ &= \left[\Sigma_{k=1}^3 a_k X_k'X_k\right]^{-1} \left[\Sigma_{k=1}^3 a_k (X_k'X_k)\beta_k\right] \end{aligned} \tag{14}$$

where the last equality follows from the definition of β_k i.e., $X_k'Y_k = (X_k'X_k)\beta_k$.

Since $e_k(\beta_k) = Y_k - X_k\beta_k = 0$ we have that $e_k(b) = (Y_k - X_kb) - (Y_k - X_k\beta_k) = X_k(\beta_k - b)$. Hence

$$e_k(b)'e_k(b) = (\beta_k - b)'X_k'X_k(\beta_k - b).$$

That is, the sum of squared errors for equation k is an ellipse in $(b_1,...,b_N)$-space and the ellipse is centred on β_k.[31] Let $R_k^2(b)$ be the goodness of fit for equation k when the slope is b. Define $A_k = X_k'X_k/Y_k'Y_k$. Then

$$R_k^2(b) \equiv 1 - e_k'(b)e_k(b)/Y_k'Y_k = 1 - (\beta_k - b)'A_k(\beta_k - b).$$

$R_k^2(b)$ is thus also an ellipse centred on β_k. It is displayed in figure 9 although it obviously need not be a circle. $R_k^2(b) = 1$ at the ellipse's centre $b = \beta_k$ and the size of the level curves $R_k^2(b) = c$ grow as c shrinks.

Consider the left panel of figure 10. It draws the level curves for the ellipses $R_k^2(b)$ and $R_{k'}^2(b)$. It also plots the tangencies between the ellipses. A typical tangency is a point b that solves

$$\max_b R_k^2(b) \text{ subject to } R_{k'}^2(b) = c$$

for some constant c. As is well know, the solution to this problem is described by $b(a)$ of equation (14) in the case where there are two equations.[32] [33]

---

[31]Subscripts always refer to equations ($k = 1,2,3$) except in the case of b where subscripts refer to elements of the vector $b = (b_1,...,b_N)$.

[32]Let $\pounds = 1 - (Y_k - X_kb)'(Y_k - X_kb)/Y_k'Y_k - \lambda\left[c - 1 + (Y_j - X_jb)'(Y_j - X_jb)/Y_j'Y_j\right]$. The first-order condition is

$$\partial\pounds/\partial b = 2(X_k'X_k + \lambda X_j'X_j)b - 2(X_k'Y_k + \lambda X_j'Y_j) = 0.$$

Letting $\lambda = a_j/a_k$ for some pair (a_k,a_j), multiplying through by $a_k/2$ and reorganizing yields

$$b = \left[a_kX_k'X_k + a_jX_j'X_j\right]^{-1}\left(a_kX_k'Y_k + a_jX_j'Y_j\right).$$

The generalization to two constraints yields the expression for $b(a)$ in equation (14).

[33]There is a slight abuse of notation here. The correct notation is $a = (a_1,a_2,a_3)$, $b(a)$ and $b((a_1,a_2,a_3))$. We drop one of the sets of parentheses and write $b(a_1,a_2,a_3)$.

Figure 9: An $R_k^2(b)$ Ellipse

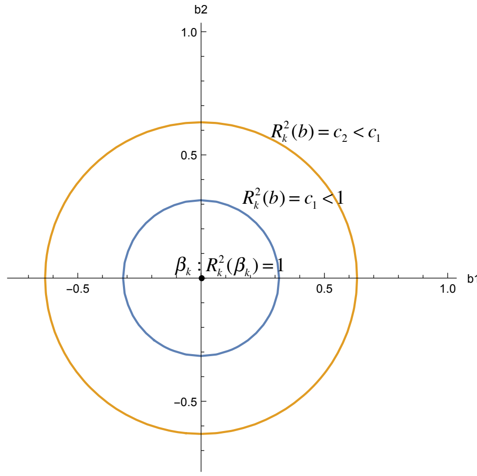
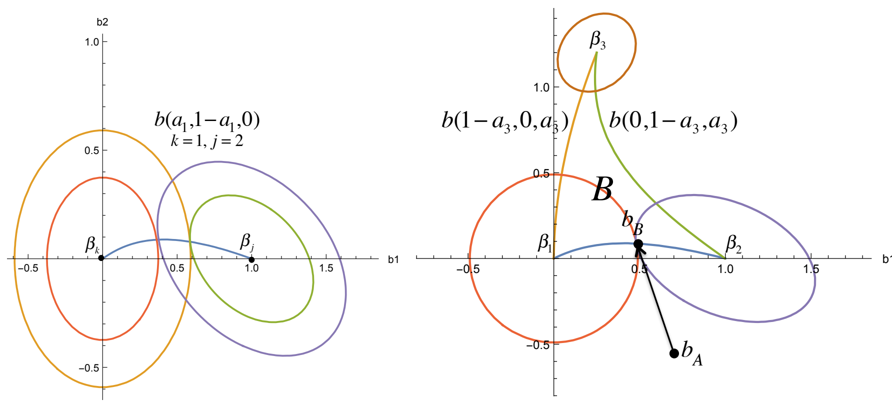


Figure 10: Set of Tangencies Between Ellipses k and j



Consider the set of tangencies or *décolletage* between the ellipses for $k = 1$ and $k' = 2$. The décolletage is displayed in the left panel of figure 10. Its equation is given by $b(a_1,1 - a_1,0)$ as a_1 varies from 0 to 1. Likewise, the tangency curve between ellipses $k = 1$ and $k' = 3$ is given by $b(a_1,0,1 - a_1)$ for $a_1 \in [0,1]$ and between ellipses $k = 2$ and $k' = 3$ is given by $b(0,1 - a_3,a_3)$ for $a_3 \in [0,1]$. All three tangency curves are displayed in the right panel of figure 10. Let B be the area bounded by these tangency curves.

Theorem 2 *Consider the problem*

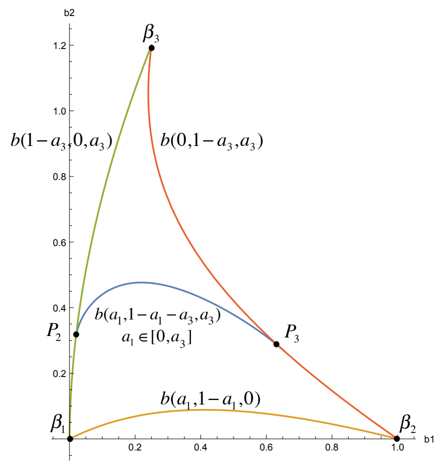
$$\max_b \left\{ \min_k \left\{ R_k^2(b) \right\} \right\} = \max_b \left\{ \min \left\{ R_1^2, R_2^2, R_3^2 \right\} \right\} \qquad \text{(Program P)}$$

and let b^* be a solution.

a. b^* exists and is unique

b. The solution b^* is a matrix-weighted average of the π_f^V, π_f^W and π_f^T and has the form $b(a^*) = \left[\Sigma_{k=1}^3 a_k^* X_k' X_k\right]^{-1} \left[\Sigma_{k=1}^3 a_k^* (X_k' X_k) \beta_k\right]$ for some unique $a^* \in \Delta$.

Figure 11: The Set B of 'Efficient' b



c. *The two worst-fitting equations have equal* R^2*s.*

d. $\widehat{\pi}_f$ *is Pareto optimal: Any* π_f *that improves the* R^2 *of one equation lowers the* R^2 *of at least one of the two remaining equations.*

Proof of part (a): Since B is compact (closed and bounded in R^N), Program P must have a solution on B. To see that the solution is unique start considering an arbitrary point b_A outside of B. b_A is displayed in the right panel of figure 10. For such a point, the move towards B indicated by the arrow must increase R_1^2 because it moves b_A to a tighter level curve for ellipse 1. Likewise, the move raises R_2^2 and R_3^2. It follows that the solution to our problem cannot be a point like b_A. Restated, any solution b^* must lie in B. Now consider a solution $b_A \in B$. Since b_A is a solution it must be that $R_1^2(b_A) = R_2^2(b_A) = R_3^2(b_A)$. Consider another candidate solution b_B. By part (b), $b_B \in B$. Comparing any two points in B, one must have a lower R_k^2 for some k and a higher R_j^2 for some $j \neq k$. See figure 10. Thus, comparing b_A and b_B in B, it must be that $R_j^2(b_B) > R_j^2(b_A) = R_k^2(b_A) > R_k^2(b_B)$ for appropriate choice of k and j. It follows that $R_j^2(b_B) \neq R_k^2(b_B)$ i.e., b_B cannot be a solution.

Proof of part (b): We begin by showing that each point in B can be represented as $b(a)$ of equation (14) for some unique $a = (a_1,1-a_1-a_3,a_3) \in \Delta$. Fix a_3 and let a_1 vary on the unit simplex from 0 to $1-a_3$. When $a_1 = 0$, $b(a_1,1-a_1-a_3,a_3) = b(0,1-a_3,a_3)$ is a point on the tangency curve connecting ellipses 2 and 3. See point P_3 in figure 11. When $a_1 = 1-a_3$, $b(a_1,1-a_1-a_3,a_3) = b(1-a_3,0,a_3)$ is a point on the tangency curve connecting ellipses 1 and 3. See point P_2 in figure 11. It follows from continuity that the curve or set

$$\widetilde{b}(a_3) = \{b : b = b(a_1,1-a_1-a_3,a_3), a_1 \in [0,a_3]\}$$

is a curve between points P_2 and P_3. See figure 11.

So far we have fixed a_3. We now vary a_3. When $a_3 = 0$, this curve coincides with the tangency curve connecting ellipses 1 and 2. When $a_3 = 1$ the curve is degenerate at the point β_3. It follows that as we vary a_3 from 0 to 1, the curve $\widetilde{b}(a_3)$ crosses over every point in B. This establishes that each point $b \in B$ can be represented by a point $b(a)$ for some $a \in \Delta$ and each point $a \in \Delta$ generates a point $b(a) \in B$. Since the unique solution to Problem P lies in B it can be represented by $b(a^*)$ for some unique $a^* \in \Delta$.

Proof of part (c): There is one condition under which the three R^2 statistics will not be the same as suggested in the proof to part (a). Suppose that when $a_k = 0$, $R_k^2(b) > max\{min\{R_{k'}^2, R_{k''}^2\}\}$. In this case, adding progressively more weight onto equation k will (weakly) improve its fit and (weakly) decrease the fit of the remaining equations such that $R_k^2(b) = R_{k'}^2(b) = R_{k'}^2(b)$ is not possible.

Proof of part (d): See the proof of part (a).

Algorithm

1. Compute $X_k' X_k$, $A_k = (X_k' X_k)/Y_k' Y_k$ and $\beta_k = (X_k' X_k)^{-1} X_k' Y_k$, $k = 1,2,3$. Do this outside of the grid search.

2. do $a_1 = 0,...,1$; do $a_2 = 0,...,1 - a_1$.

   (a) Set $a_3 = 1 - a_1 - a_2$.

   (b) Calculate $b(a) = \left[\Sigma_{k=1}^3 a_k^2 X_k' X_k\right]^{-1} \left[\Sigma_{k=1}^3 a_k^2 (X_k' X_k)\beta_k\right]$.

   (c) Calculate $R_k^2(b(a)) = 1 - (\beta_k - b(a))' A_k (\beta_k - b(a))$.

   (d) Calculate the 3 slope parameters from regressions of fitted on actual. Call these $Slope_k(a)$.

   (e) end the two do loops.

3. For each (a_1,a_2), calculate $R^2(a_1,a_2) = \min_k \left[R_k^2(b(a))\right]$.

4. Calculate $\min_{(a_1,a_2)} R^2(a_1,a_2)$.

Appendix D. Proof of Theorem 3

We now prove that σ and π_{fi} are not identified given our current framework. A sketch is as follows: suppose that we know the value of σ with certainty and wish to calculate the vector of productivity terms $\{\pi_{fi}\}$ consistent with the system. This can be done by minimizing a sum of least squared deviations that will equal zero if the system fits exactly and will be minimized if the system is over-identified. We then impose a different value of σ, calculate the resulting $\{\pi_{fi}\}$ that minimize the distance, and recalculate the sum of squared deviations. If the sum of squared deviations does not change, σ is not identified. We show that this is the case. Starting with equations V'-T', start by assuming that w_{fi}, σ and $(\delta_{gi}/\delta_{g,US})^{1-\sigma}$ are known along with all production, employment and endowment data. We introduce the following terms for bevity

$$H_{fi} \equiv \left(\sum_{g=1}^{G} \frac{d_{fgi'} Q_{gi}}{V_{fi}} (\delta_{gi})^{\sigma-1} \right) \qquad \tilde{V}_{fi} = \delta_{fi}' V_{fi}.$$

Define the following $3Nx1$ vector W_f, the $Nx1$ vector Π_f, the $3Nx1$ vector μ_f, and the $3NxN$ matrix $Z_f(\sigma)$ as below.

$$W_f = \begin{bmatrix} F_{f,1} \\ F_{f,2} \\ \vdots \\ F_{fN} \\ \bar{d}_{f,1} \\ \bar{d}_{f,2} \\ \vdots \\ \bar{d}_{fN} \\ H_{f,1} \\ H_{f,2} \\ \vdots \\ H_{fN} \end{bmatrix}, \Pi_f = \begin{bmatrix} (\pi_{f,1})^{1-\sigma} \\ (\pi_{f,2})^{1-\sigma} \\ \vdots \\ (\pi_{fN})^{1-\sigma} \end{bmatrix}, \mu_f = \begin{bmatrix} \mu_{f,1} \\ \mu_{f,2} \\ \mu_{f,3} \\ \vdots \\ \mu_{f,3N} \end{bmatrix},$$

$$Z_f(\sigma) = \begin{bmatrix} (1-s_1)(w_{f1})^{\sigma}\tilde{V}_{f,1} & -s_1(w_{f,2})^{\sigma}\tilde{V}_{f,2} & \dots & -s_1(w_{fN})^{\sigma}\tilde{V}_{fN} \\ -s_2(w_{f,1})^{\sigma}\tilde{V}_{f,1} & (1-s_2)(w_{f,2})^{\sigma}\tilde{V}_{f,2} & \dots & -s_2(w_{fN})^{\sigma}\tilde{V}_{fN} \\ \vdots & \vdots & \ddots & \vdots \\ -s_N(w_{f,1})^{\sigma}\tilde{V}_{f,1} & -s_N(w_{f,2})^{\sigma}\tilde{V}_{f2} & \dots & (1-s_N)(w_{fN})^{\sigma}\tilde{V}_{fN} \\ \delta'_{f,1}(w_{f,1})^{\sigma} & 0 & 0 & 0 \\ \delta'_{f,2}(w_{f,2})^{\sigma} & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \delta'_{fN}(w_{fN})^{\sigma} \\ (w_{f,1})^{\sigma} & 0 & 0 & 0 \\ 0 & (w_{f,2})^{\sigma} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & (w_{fN})^{\sigma} \end{bmatrix}$$

In all that follows, assume that all matrices are non-singular. Using these, the system can be written as $M_f = Z_f \Pi_f + \mu_f$. Notice that $Z_f(\sigma_1)G(\sigma_1,\sigma_2) = Z_f(\sigma_2)$, where $G(\sigma_1,\sigma_2)$ is a NxN diagonal matrix

$$G(\sigma_1,\sigma_2) = \begin{bmatrix} (w_{f,1})^{\sigma_2-\sigma_1} & 0 & \dots & 0 \\ 0 & (w_{f,2})^{\sigma_2-\sigma_1} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & (w_{fN})^{\sigma_2-\sigma_1} \end{bmatrix}.$$

The vector of deviations can be written as $\mu_f = M_f - Z_f \Pi_f$. Therefore the Π_f that provides the best fit is $\Pi_f^{\dagger} \equiv \left[Z_f' Z_f\right]^{-1} Z_f' M_f$. The smallest set of deviations can therefore be written as $\mu_f^{\dagger} = \left[I_N - Z_f \left[Z_f' Z_f\right]^{-1} Z_f'\right] W_f$ where I_N is an NxN identify matrix. Because the identify matrix and W_f are invariant to σ, the vector of deviations will change if and only if the projection matrix $Z_f \left[Z_f' Z_f\right]^{-1} Z_f'$ changes where Z_f' is the transpose of Z_f. We now show that the projection matrix is invariant to σ. For any matrix $Z_f(\sigma)$, define the projection matrix for $Z_f(\sigma)$ as

$$P(\sigma) \equiv Z_f(\sigma)\left[Z_f(\sigma)' Z_f(\sigma)\right]^{-1} Z_f(\sigma)'.$$

Consequently, we can write the projection matrix for $Z_f(\sigma_2)$ as:

$$P(\sigma_2) = Z_f(\sigma_2)\left[Z_f(\sigma_2)' Z_f(\sigma_2)\right]^{-1} Z_f(\sigma_2)'.$$

Noting that $Z_f(\sigma_2) = Z_f(\sigma_1)G(\sigma_1,\sigma_2)$, $P(\sigma_2)$ can be rewritten as

$$P(\sigma_2) = Z_f(\sigma_1)G(\sigma_1,\sigma_2)\left[(Z_f(\sigma_1)G(\sigma_1,\sigma_2))' Z_f(\sigma_1)G(\sigma_1,\sigma_2)\right]^{-1}(Z_f(\sigma_1)G(\sigma_1,\sigma_2))'.$$

Using the fact that the inverse of a product of matrices is equal to the product of the inverses, along with the fact that the transpose of a diagonal matrix is equal to itself, we can write:

$$P(\sigma_2) = Z_f(\sigma_1)G(\sigma_1,\sigma_2)G(\sigma_1,\sigma_2)^{-1}\left[Z_f(\sigma_1)'Z_f(\sigma_1)\right]^{-1}G(\sigma_1,\sigma_2)^{-1}G(\sigma_1,\sigma_2)Z_f(\sigma_1)'.$$

$$P(\sigma_2) = Z_f(\sigma_1)\left[Z_f(\sigma_1)'Z_f(\sigma_1)\right]^{-1}Z_f(\sigma_1)' = P(\sigma_1)$$

Therefore, the projection matrix of $Z_f(\sigma)$ is invariant to σ and, therefore, vector of residuals in invariant to σ as is the sum of squared residuals and σ is not identified. QED.

Appendix E. Consumption Similarity Calculation

We now show how to calculate consumption similarity adjusted net exports. Matrices and vectors are presented in bold and scalars in normal font. We adjust each of the vectors $\mathbf{X_i}$ and $\mathbf{M_{ij}}$ to be consistent with consumption similarity. To do so, start with the definitions

$$\mathbf{X_i} \equiv \mathbf{Q_i} - \mathbf{B_{ii}}\mathbf{Q_i} - \mathbf{C_{ii}}, \tag{15}$$

$$\mathbf{M_{ij}} \equiv \mathbf{B_{ji}}\mathbf{Q_i} + \mathbf{C_{ij}}. \tag{16}$$

Following the definition of consumption similarity ($c_{gij} = s_i c_{g,world,j}$), we have

$$\mathbf{C_{ij}} = s_i \mathbf{C_{Wj}}. \tag{17}$$

In addition, world goods market clearing requires

$$\mathbf{Q_i} = \mathbf{C_{Wi}} + \sum_{j=1}^{N} \mathbf{B_{ij}}\mathbf{Q_j}. \tag{18}$$

Using equations (15)-(18), we define the vectors of exports and imports that are consistent with consumption similarity as follows

$$\mathbf{X_i^{CS}} \equiv \left[\mathbf{I_G} - \mathbf{B_{ii}}\right]\mathbf{Q_i} - s_i\left[\mathbf{Q_i} - \sum_{j=1}^{N}\mathbf{B_{ij}}\mathbf{Q_j}\right],$$

$$\mathbf{M_{ij}^{CS}} \equiv \mathbf{B_{ji}}\mathbf{Q_i} + s_i\left[\mathbf{Q_j} - \sum_{i=1}^{N}\mathbf{B_{ji}}\mathbf{Q_i}\right],$$

where $\mathbf{I_G}$ is a GxG identity matrix. To impose consumption similarity in the g^{th} industry, we replace the g^{th} rows of $\mathbf{X_i}$ and $\mathbf{M_{ij}}$, respectively, with the g^{th} rows of $\mathbf{X_i^{CS}}$ and $\mathbf{M_{ij}^{CS}}$, respectively. These are then used to create the consumption similarity consistent trade matrix $\mathbf{T^{CS}}$ (analogous $\mathbf{T}$) and the factor content of trade $\mathbf{F}^{*\mathbf{CS}} \equiv \mathbf{D_f^*}(\mathbf{I_{NG}} - \mathbf{B})^{-1}\mathbf{T^{CS}}$. Note that consumption similarity involves only affecting consumption patterns and not the use of intermediate inputs in production $\mathbf{B}$.

References

Antràs, Pol, "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution," *The B.E. Journal of Macroeconomics*, April 2004, *4* (1), 1–36.

Antweiler, Werner and Daniel Trefler, "Increasing Returns and All That: A View From Trade," *American Economic Review*, March 2002, *92* (1), 93–119.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

Banerjee, Abhijit V. and Esther Duflo, "Growth Theory through the Lens of Development Economics," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1 of *Handbook of Economic Growth*, Elsevier, June 2005, chapter 7, pp. 473–552.

Barro, Robert J. and Jong Wha Lee, "A new data set of educational attainment in the world, 1950Ð2010," *Journal of Development Economics*, September 2013, *104* (C), 184–198.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott, "Comparative Advantage and Heterogeneous Firms Comparative Advantage and Heterogeneous Firms," *Review of Economic Studies*, January 2007, *74* (1), 31–66.

Burstein, Ariel and Jonathan Vogel, "Factor Prices and International Trade: A Unifying Perspective," NBER Working Papers 16904, National Bureau of Economic Research, Inc March 2011.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1 of *Handbook of Economic Growth*, Elsevier, June 2005, chapter 9, pp. 679–741.

_ **and James Feyrer**, "The Marginal Product of Capital," *Quarterly Journal of Economics*, May 2007, *122* (2), 535–568.

_ **and Wilbur John Coleman**, "The World Technology Frontier," *American Economic Review*, June 2006, *96* (3), 499–522.

Ciccone, Antonio and Giovanni Peri, "Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990," *The Review of Economics and Statistics*, November 2005, *87* (4), 652–663.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Elhanan Helpman, Kenneth Rogoff, and Gita Gopinath, eds., *Handbook of International Economics*, Vol. 4, Elsevier, 2014, pp. 197–261.

Davis, Donald R. and David E. Weinstein, "An Account of Global Factor Trade," *American Economic Review*, December 2001, *91* (5), 1423–1453.

Diamond, Peter, Daniel McFadden, and Miguel Rodriguez, "Measurement of the Elasticity of Factor Substitution and Bias of Technical Change," in Melvyn Fuss and Daniel L. McFadden, eds., *Production Economics: A Dual Approach to Theory and Applications Volume II: Applications of the Theory of Production*, Amsterdam: North-Holland, 1978, pp. 125–147.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *75* (5), 1741–1779.

Fadinger, Harald, "Productivity differences in an interdependent world," *Journal of International Economics*, July 2011, *84* (2), 221–232.

Fally, James R. Markusen Justin Caron Thibault, "International trade puzzles: A solution linking production and preferences," *Quarterly Journal of Economics*, August 2014, *129* (3), 1501–1552.

Feenstra, Robert C. and Gordon H. Hanson, "Aggregation Bias in the Factor Content of Trade: Evidence from U.S. Manufacturing," *American Economic Review*, May 2000, *90* (2), 155–160.

Ferguson, Thomas S., *Mathematical Statistics: A Decision Theoretic Approach*, New York: Academic Press, 1967.

Fitzgerald, Doireann and Juan Carlos Hallak, "Specialization, factor accumulation and development," *Journal of International Economics*, December 2004, *64* (2), 277–302.

Freeman, Richard and Remco H. Oostendorp, "The Occupational Wages around the World (OWW) Database: Update for 1983-2008," World Development Report 2013 Background Paper No., World Bank, 2012.

Gabaix, Xavier, "The Factor Content of Trade: A Rejection of the Heckscher-Ohlin-Leontief Hypothesis," 1997. Mimeo, Harvard University.

_ , "The Factor Content of Trade: A Rejection of the Heckscher-Ohlin-Leontief Hypothesis," 1997. Mimeo, Harvard University.

Gollin, Douglas, David Lagakos, and Michael E. Waugh, "Agricultural Productivity Differences across Countries," *American Economic Review*, May 2014, *104* (5), 165–70.

Hallak, Juan Carlos, "Product Quality and the Direction of Trade," *Journal of International Economics*, January 2006, *68* (1), 238–65.

Hunter, Linda C. and James R. Markusen, "Per-Capita Income as a Determinant of Trade," in Robert C. Feenstra, ed., *Empirical Methods for International Trade*, Cambridge: MIT Press, 1988, pp. 89–109.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, February 1992, *107* (1), 35–78.

Khandelwal, Amit, "The Long and Short (of) Quality Ladders," *Review of Economic Studies*, October 2010, *77* (4), 1450–1476.

Krugman, Paul R., "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, December 1980, *70* (5), 950–959.

Lagakos, David and Michael E. Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, April 2013, *103* (2), 948–80.

Leamer, Edward E., *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, New York, MA: John Wiley & Sons Inc., 1978.

_ , "The Leontief Paradox Reconsidered," *Journal of Political Economy*, 1980, *88*, 495–503.

Levchenko, Andrei A. and Jing Zhang, "Ricardian productivity differences and the gains from trade," *European Economic Review*, January 2014, *65* (C), 45–65.

Marshall, Kathryn G., "International productivity and factor price comparisons," *Journal of International Economics*, 2012, *87* (2), 386–390.

Maskus, Keith E. and Shuichiro Nishioka, "Development-Related Biases in Factor Productivities and the HOV Model of Trade," *Canadian Journal of Economics*, May 2009, *42* (2), 519–553.

Nishioka, Shuichiro, "International differences in production techniques: Implications for the factor content of trade," *Journal of International Economics*, 2012, *87* (1), 98–104.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," NBER Working Papers 20452, National Bureau of Economic Research, Inc September 2014.

Parro, Fernando, "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 72–117.

Puzzello, Laura, "A proportionality assumption and measurement biases in the factor content of trade," *Journal of International Economics*, May 2012, *87* (1), 105–111.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu, "Agriculture and aggregate productivity: A quantitative cross-country analysis," *Journal of Monetary Economics*, March 2008, *55* (2), 234–250.

Schott, Peter, "One Size Fits All? Heckscher-Ohlin Specialization in Global Production," *American Economic Review*, June 2003, *93* (2), 686–708.

_ , "Across-Product versus Within-Product Specialization in International Trade," *Quarterly Journal of Economics*, May 2004, *119* (2), 647–678.

Sutton, John and Daniel Trefler, "Capabilities, Wealth and Trade," *Journal of Political Economy*, forthcoming.

Trefler, Daniel, "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy," *Journal of Political Economy*, February 1993, *101* (1), 138–160.

_ , "The Case of the Missing Trade and Other Mysteries," *American Economic Review*, December 1995, *85* (5), 1029–1046.

_ **and Susan Chun Zhu**, "Beyond the Algebra of Explanation: HOV for the Technology Age," *American Economic Review Papers and Proceedings*, May 2000, *90* (2), 145–149.

_ **and** _ , "The structure of factor content predictions," *Journal of International Economics*, November 2010, *82* (2), 195–207.

Vanek, Jaroslav, "The Factor Proportions Theory: The N-Factor Case," *Kyklos*, October 1968, *21*, 749–56.

Waugh, Michael E., "International Trade and Income Differences," *American Economic Review*, December 2010, *100* (5), 2093–2124.