NBER WORKING PAPER SERIES

EXPORT MARKETS AND LABOR ALLOCATION IN A LOW-INCOME COUNTRY

Brian McCaig
Nina Pavcnik

Working Paper 20455
http://www.nber.org/papers/w20455

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2014, Revised January 2018

We thank Hanh Nyugen and Truong Thi Thu Trang for excellent research assistance and Eric Edmonds, Teresa Fort, Doug Irwin, Amit Khandelwal, discussants and seminar participants at several universities and conferences for comments. We thank the International Food Policy Research Institute (IFPRI) and the World Bank for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w20455.ack

Export Markets and Labor Allocation in a Low-income Country
Brian McCaig and Nina Pavcnik
NBER Working Paper No. 20455
September 2014, Revised January 2018
JEL No. F14,F16,O17,O47

ABSTRACT

We study the effects of a positive export shock on labor allocation between the informal, microenterprise sector and the formal firm sector in a low-income country. The U.S.-Vietnam Bilateral Trade Agreement led to large reductions in U.S. tariffs on Vietnamese exports. We find that the share of manufacturing workers in Vietnam in the formal sector increased by 5 percentage points in response to the U.S. tariff reductions. The reallocation was greater for workers in more internationally integrated provinces and for younger cohorts. We estimate the gap in labor productivity within manufacturing across the informal and formal sectors. This gap and the aggregate labor productivity gain from the export-induced reallocation of workers across the two sectors are reduced when we account for worker heterogeneity, measurement error, and differences in labor intensity of production.

Brian McCaig
Department of Economics,
Wilfrid Laurier University
75 University Avenue West
Waterloo, ON, N2L 3C5 Canada
bmccaig@wlu.ca

Nina Pavcnik
Department of Economics
6106 Rockefeller Center
Dartmouth College
Hanover, NH 03755
and NBER
nina.pavcnik@dartmouth.edu

1. Introduction

Low- and middle-income countries have dramatically increased their participation in world trade over the last two decades, with their share of total world exports growing from 21 to 43 percent and export growth outpacing output growth in these economies between 1992 and 2008 (Hanson (2012)). Within low-income countries, 70-80 percent of employment is in informal, household-run microenterprises, which are substantially less productive than their formal-sector counterparts.[1] Studies document a systematic negative relationship between the prevalence of informal microenterprise employment and economic development. These studies attribute aggregate income differences across countries to the inefficient allocation of inputs across sectors and firms.[2] As a result, the reallocation of workers from microenterprises to more productive establishments in response to rising exports could increase aggregate productivity.

This paper examines the relationship between international trade and the allocation of labor across the informal microenterprise and formal sectors in a low-income country setting. International trade can contribute to economic development if it promotes the reallocation of workers out of informal microenterprises toward relatively more productive establishments. However, exploring the role of this channel has been challenging for several reasons. First, data constraints often make it difficult to measure such reallocation. While literature has examined how trade affects labor allocation in less developed countries (see Goldberg and Pavcnik (2007) and Harrison, McLaren, and McMillan (2011) for surveys), data availability has limited most of this work to employment in the formal sector or in urban areas.[3] In addition, the data in these studies rarely distinguish between work in informal microenterprises and formal firms.[4] Second, employment in

[1] See Gollin (2002, 2008), Tybout (2000, 2014), and Nataraj (2011) for estimates of high rates of informal self-employment and La Porta and Shleifer (2008, 2014) and Nataraj (2011) for large productivity gaps.

[2] See Gollin (2002, 2008) and La Porta and Shleifer (2008, 2014) for evidence on the relationship between informal, microenterprise employment and aggregate development and Banerjee and Duflo (2005), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and McMillan and Rodrik (2011) for the relationship between resource allocation and development.

[3] See Revenga (1997), Currie and Harrison (1997), Wacziarg and Wallack (2004), and Topalova (2010) for studies on trade policy and employment allocation across industries or firms using data based on formal manufacturing firms above a certain employment cut-off; Feliciano (2001) and Attanasio, Goldberg, and Pavcnik (2004) for studies using household surveys representative of labor markets in major urban areas; and Menezes-Filho and Muendler (2011) and Dix-Carneiro (2014) for studies using administrative employee-employer data that cover formal employers. Topalova (2010), McCaig (2011), and Kovak (2013) use nationally representative household surveys, but they do not study reallocation across informal microenterprises and formal firms.

[4] Goldberg and Pavcnik (2003), Bosch, Goñi-Pacchioni, and Maloney (2012), Paz (2014), and Menezes-Filho and Muendler (2011) examine the impact of trade on (formal) firm compliance with labor legislation (i.e. severance payments or social security legislation) for a worker. Their definition of informality is worker based.

the formal enterprise sector and international trade both tend to expand during the process of growth and urbanization in low-income countries, making it difficult to identify the causal effect of international trade on worker allocation across employers. Finally, the potential aggregate labor productivity gains from labor reallocation through this channel depend on the size of the labor productivity gap between informal microenterprises and formal firms. This gap is difficult to measure because comprehensive data on informal microenterprises is scarce, and because informal microenterprises and formal firms differ in dimensions other than productivity, including the composition of the workforce and hours worked, record-keeping standards, and labor intensity of production. For these reasons, the literature is missing a potentially important dimension of labor reallocation through which trade affects economic development.

The paper addresses the above challenges by focusing on Vietnam, which has nationally representative household surveys that includes the informal microenterprise sector, and which was subjected to a large, plausibly exogenous (but positive) export shock with the 2001 U.S.-Vietnam Bilateral Trade Agreement (henceforth, the BTA).

In low-income countries, nationally representative labor force data, covering workers in all types of employers, is more commonly available than firm-level data that captures the entire firm distribution. We use labor force data to analyze trade-induced shifts of labor across typically unobserved points in the firm distribution, i.e., between informal and formal firms. Specifically, we use labor force data from several Vietnam Household Living Standards Surveys (VHLSS), which cover workers in all industries and types of employers and record whether a worker works for an employer in the household business (informal) sector or the registered enterprise sector.[5] Our definition of work in an informal microenterprise is based on the registration status of the business. It is consistent with the definition of informal firms in other low-income countries and has a direct connection to studies on the role of microenterprises in economic development. In Vietnam, household businesses employed about 85% of workers economy-wide, and 66% in the manufacturing sector, at the onset of the trade agreement.

We combine the labor force data with large and plausibly exogenous variation in declines in export costs induced by the U.S.-Vietnam BTA. The principle trade policy change in the BTA was a significant, immediate drop in U.S. tariffs on Vietnamese exports. This drop averaged 20.9

[5] A household business is not registered as an enterprise under Vietnam's Enterprise Law. Not all businesses are required to register. See Section 4 for details on the definition and how it compares to other settings.

percentage points, which substantially lowered the cost of exporting Vietnamese products to the United States. This resulted in a substantial shock to Vietnam's trade. Between 2001 and 2004, exports to the United States grew from 1.1 to 5.0 billion USD (Figure 1), from 7.1 to 19.0 percent of total exports (Figure 2), and from 3.6 to 10.4 percent of Vietnam's GDP.

Our empirical setting overcomes a key challenge in identifying the effect of exporting on the allocation of workers by using heterogeneity in policy-driven export cost reductions across industries: Vietnamese industries that faced greater declines in U.S. tariffs observed greater export growth (Figure 3). The agreement lowered industry-specific tariffs when the United States moved Vietnam from the pre-existing Column 2 to the pre-existing MFN U.S. tariff schedule, rather than by industry-specific contemporaneous negotiations over tariff lines (McCaig (2011)). This means that the industry-specific declines in U.S. tariffs were plausibly exogenous and not precipitated by industry-specific economic conditions in Vietnam during the early 2000s (see Section 3 for details). Importantly, tariff changes are not spuriously correlated with pre-existing or concurrent global demand or supply shocks to Vietnamese products that occur at the same time in industries with greater declines in U.S. tariffs. While Vietnamese industries that faced larger declines in U.S. tariffs observed greater export growth (driven by U.S. export expansion), we show that U.S tariff declines were not predictive of Vietnamese export growth prior to the agreement, nor were they correlated with Vietnamese export growth to the European Union, another high-income export destination.

After examining the effect of export costs on the movement of employment from the household business to the enterprise sector, we assess the potential gains in aggregate labor productivity from this trade-induced shift of workers across the two sectors. By combining comprehensive data on informal microenterprises from the household business module of the VHLSS, with the census of registered enterprises, we provide one of the first estimates of the labor productivity gap between the informal and formal sectors for a low-income country. Conceptually, we follow the approach in the macroeconomic development literature (see Caselli (2005) for a survey). We adjust the labor productivity gap for differences in worker composition and hours worked across the two sectors using worker-level information from the VHLSS.[6] We also assess the

[6] Our approach relates to the literature on the labor productivity gap between agriculture and non-agriculture. See Gollin, Lagakos, and Waugh (2014), Vollrath (2014) and Herrendorf and Schoellman (2015) for measurement and Caselli (2005) and Restuccia, Yang, and Zhu (2008) for the role of these labor productivity gaps in explaining aggregate productivity differences across countries.

potential role of differences in measurement error and the output elasticity of labor across the two sectors for the size of the labor productivity gap.

The reallocation of labor from household businesses to employers in the enterprise sector provides an important margin of adjustment to exporting. Approximately half of the aggregate decline in the share of workers in household businesses during the early 2000s derives from such reallocation within industries.[7] The within-industry component is particularly pronounced in manufacturing, which experienced the largest tariff cuts and greatest export growth. Importantly, industries that experience larger declines in tariffs on Vietnamese exports to the U.S. experienced greater decreases in the share of household business employment. The estimated magnitudes for manufacturing imply that export opportunities due to the U.S. tariff reductions increased the share of employment in the enterprise sector in manufacturing by 4.9 percentage points: about 207,000 workers. Our estimates reflect short run responses and may underestimate the long run effects of the BTA on employment in the formal sector, as the economy has more time to adjust. We perform several robustness and falsification checks. We find no effects of the BTA on household business employment prior to its implementation. Additionally, the results are robust to using self-employment as a measure of informality, a measure potentially more comparable across countries than one based on the country-specific legal definition. Moreover, the results are robust to controlling for the sorting of workers across sectors based on observable and time-invariant unobservable worker characteristics.

Our results are consistent with models that predict a reallocation of workers away from self-employment into wage employment and towards larger, more productive firms in response to shocks that raise aggregate wages. For example, Melitz (2003) style models suggest that a reduction in export market tariffs leads to a reallocation of labor toward more productive firms as the wage rises (Demidova and Rodriguez-Clare (2013)). More generally, Lucas (1978) style models, such as Gollin (2008), predict a reallocation of individuals from self-employment toward wage work in firms run by managers with greater managerial talent as the aggregate wage rises. This reallocation mechanism is further supported by evidence of wage increases and poverty reductions in McCaig (2011) and Fukase (2013), where BTA-induced declines in US tariffs are associated with greater increases in wages (especially for less educated workers) and decreases in poverty in areas of

[7] The remaining half of the decline derives from the relative contraction of industries that tend to concentrate production in microenterprises, namely agriculture and aquaculture.

Vietnam more exposed to exporting. Not all individuals are affected equally by exporting opportunities. Younger workers and workers in more internationally integrated provinces are more likely to reallocate, which is consistent with lower adjustment costs to trade shocks among the young and those with lower geographic mobility costs (Dix-Carneiro (2014), Hanson (1996)).

Our study also relates to the literature on the effects of trade policy on the allocation of labor across industries, which finds limited net industry employment adjustment in response to tariff declines in the short run in less developed countries.[8] In Section 5.5, we also do not find shifts in the structure of total industry employment with declines in export costs. Earlier studies primarily examined the domestic market consequences of unilateral import liberalizations. We focus on export market liberalization and our findings of employment shifting toward employers in the formal sector are consistent with theory predictions.[9]

More generally, our analysis highlights the role of output-market factors (i.e. demand-side constraints) in influencing the allocation of resources between informal microenterprises and formal firms. Many studies focus on the effects of the removal of input-market distortions (i.e. supply-side constraints) on the growth and formalization of microenterprises (see surveys by Banerjee (2013) and Banerjee, Karlan and Zinman (2015) for microcredit; Bruhn and McKenzie (2014) for business registration; and McKenzie and Woodruff (2014) for business training). Our study complements this literature by focusing on an output-market (i.e. product demand-side) policy change that disproportionally benefits and expands better performing firms (Melitz (2003)) and thereby reallocates employment away from microenterprises toward formal firms.

We contribute to the recent literature on the role of resource allocation across heterogeneous firms and sectors in aggregate productivity differences across countries.[10] The estimates of the aggregate labor productivity gains from the BTA-induced reallocation from the informal to the formal sector depend critically on the estimates of the labor productivity gap. The usual measure, based on average revenue product of labor, suggests a gap of 9. We show that it is crucial to account for differences in worker composition and hours worked across the household

[8] See Goldberg and Pavcnik (2007) for a survey. Section 5.5 provides detailed discussion.
[9] See also Verhoogen (2008), Bustos (2011a, b), and Brambilla, Lederman, and Porto (2012) for the effects of exporting on firm technology, quality or skill upgrading in formal enterprises and Brambilla, Porto, and Tarozzi (2012) on the effects of exporting on microenterprises in aquaculture.
[10] Most papers study misallocation among formal firms. Recent exceptions include Hsieh and Olken (2014) and Bento and Restuccia (2016).

business and enterprise sectors. This adjustment reduces the gap to 6, as worker heterogeneity accounts for almost 40% of the original average revenue per worker gap.[11] Additional adjustments for potential differences in measurement error in revenue and hours worked across the two sectors reduce the gap to 3.7, with a further drop to 2.5 when allowing for differences in the output elasticity of labor. The BTA-induced reallocation of labor from the informal to the formal sector increased aggregate labor productivity within manufacturing by 2.8 percent per year in the two years following the BTA based on a labor productivity gap of 3.7, but the increase would be 1.5 percent with a labor productivity gap of 2.5. Overall, our analysis in Section 6 highlights the sensitivity of the estimates of the labor productivity gap and aggregate productivity gain to worker heterogeneity, measurement error in revenue or employment, and assumptions on labor intensity of production. These issues are not unique to our setting and have implications for the literature on misallocation, which relies on the size of the productivity gap to assess the aggregate gain from the elimination of distortions.

Finally, our study relates to the literature on formal sector job creation in developing countries. For workers, working in a formal firm rather than an informal microenterprise is more than simply the distinction between working in a high- versus low-productivity firm. Working in a formal firm changes the way a worker is attached to the workforce. In Vietnam, workers in the formal sector earn higher wages, are more likely to receive non-wage/salary payments, work longer hours, and are less likely to hold multiple jobs. Hence, the reallocation of workers towards the formal sector in response to the BTA may have welfare consequences for workers. This is in line with Banerjee and Duflo (2008, 2011), who argue that stable jobs with regular hours are an important feature of the middle class in developing countries.[12]

The rest of the paper is organized as follows. Section 2 summarizes a conceptual framework. Sections 3 and 4 describe the BTA and the data. Section 5 discusses our empirical methodology and results. Section 6 estimates the labor productivity gap between the household business and enterprise sectors and assesses the aggregate labor productivity change in manufacturing from BTA-induced worker movement to the enterprise sector. Section 7 concludes.

[11] Gollin et al. (2014) attribute part of the labor productivity gap between agriculture and non-agriculture to worker heterogeneity and La Porta and Shleifer (2008) document lower levels of education among managers of informal firms relative to formal firms.

[12] See as well Atkin (2009), Jensen (2012), Javorcik (2015), and Heath and Mobarak (2015).

2. A conceptual framework

We briefly discuss why tariff reductions on exports from a low-income country (corresponding to the main trade policy change in the BTA) could affect the composition of employment between the household business and formal enterprise sectors within an industry. This discussion guides the empirical framework and analysis in Section 5.

A reduction in tariffs on exports from a low-income country will increase product demand and labor demand in the country. If firms differ in underlying profitability due to heterogeneity in marginal costs of production and face a fixed cost of exporting, the reduction in variable export costs disproportionately raises the profitability of firms with a lower marginal cost of production (Melitz (2003), Demidova and Rodriguez-Clare (2013)). Firm-specific marginal cost differences might stem from differences in entrepreneurial ability of the owner/manager (Lucas (1978), Gollin (2008)) or underlying productivity (Melitz (2003)). Household businesses differ from firms in the enterprise sector in many dimensions and exhibit substantially lower productivity, perhaps owing in part to lower education or managerial ability of owners.[13] In this setting, only initially more productive firms benefit from declines in policy-induced variable export costs because only they earn high enough variable profits from increased exports to cover the fixed cost of exporting. Declines in tariffs increase product and labor demand (and profitability) among these more productive firms, while increasing the labor costs and reducing the profitability of inefficient firms that only serve the domestic market. This is predicted to shift the composition of employment away from less productive employers (such as household businesses) toward more productive employers in the enterprise sector.[14]

This mechanism does not require that household businesses and formal enterprises compete in the product market. A framework such as Melitz (2003) assumes product-market competition among the firms, implying, in our context, that household business products are imperfect substitutes for varieties produced by firms in the enterprise sector, including exported varieties. This is clearly a strong assumption. Even if household businesses and formal enterprises do not compete in product markets, exporting could affect employment in household businesses

[13] See Gollin (2008), La Porta and Shleifer (2008, 2014), de Mel, McKenzie, and Woodruff (2013), and Nataraj (2011).

[14] Mrazova and Neary (forthcoming) show that the selection effects in Melitz style models are very robust to functional form assumptions and market structure, requiring supermodularity of the profit function in marginal production costs and market access costs (export).

through the general equilibrium effects of trade on labor demand. In fact, evidence from Vietnam suggests that exporting opportunities from the BTA raise wages (McCaig (2011), Fukase (2013)). If household businesses compete for labor with firms in the enterprise sector, which disproportionately benefit from declines in export costs (Melitz (2003)), the increased labor demand among firms in the enterprise sector increases the opportunity cost of working for a household business, resulting in a relative contraction of employment in household businesses (see also Lucas (1978), Gollin (2008)). This discussion abstracts from frictions that might impede the mobility of individuals from the household business to the enterprise sector. To the extent that such frictions exist, they dampen the reallocation in response to declines in export costs, making it more difficult to detect empirically reallocation across this margin of employment after tariff declines. Likewise, firms might face different distortions across the two sectors (see Hsieh and Klenow (2009)). This would lead to lower employment in a sector facing greater distortions than in Melitz (2003) without distortions.

Reductions in trade costs also influence the relative size of industries, as emphasized in the neoclassical trade models, and this too may influence the allocation of labor between the household business and enterprise sectors. In general, the effect of trade on the composition of aggregate employment across employers via this neoclassical channel depends on the nature of the trade liberalization and the relative prevalence of household business employment in industries subject to larger declines in trade frictions. For example, in Vietnam, production in agriculture is more prone to be organized around household businesses than the apparel industry. If the trade agreement reduces the export cost of apparel (relative to agriculture), trade shifts the structure of employment away from agriculture toward apparel, reducing the aggregate share of jobs in household businesses. Our empirical framework accounts for such compositional changes.

3. Background on the U.S.-Vietnam Bilateral Trade Agreement

In this section, we describe the U.S.-Vietnam Bilateral Trade Agreement (BTA) and highlight its key features that we utilize in our empirical methodology and identification strategy in Section 5.

The BTA was implemented on December 10, 2001.[15] The agreement led to negligible changes in Vietnam's import tariff commitments to the U.S. because Vietnam already applied Most

[15] See STAR-Vietnam (2003) and McCaig (2011) for an extensive discussion of the BTA.

Favored Nation (MFN) tariffs on U.S. imports.[16] The main trade policy change was for the U.S to immediately grant Vietnam Normal Trade Relations (NTR) or MFN access to the U.S. market. Prior to the BTA Vietnam was subject to tariffs according to Column 2 of the U.S. tariff schedule. With the BTA, Vietnam became subject to MFN tariff rates. In our analysis, we use industry-level U.S. import ad valorem equivalent tariffs applied to Vietnamese exports constructed from these two tariff schedules by McCaig (2011) as the main policy variable to measure the industry-level policy cost of accessing export markets.[17]

Our identification strategy in Section 5 relies on several features of the U.S. tariff declines. Table 1 summarizes industry tariff levels and changes overall and for broad sectors. First, the U.S. tariff cuts were large, as the BTA on average reduced tariffs by 20.9 percentage points, from 23.4 to 2.5 percent. The large magnitude of tariff cuts makes it ex ante plausible to separate the effects of changes in tariffs from confounding changes in the Vietnamese economy. Our empirical methodology in Section 5 relies on the heterogeneity of tariff declines across industries to identify the effects of lower exporting costs on labor allocation across employers. Thus, a second useful feature of the BTA is that the tariff cuts varied widely across industries. As Table 1 suggests, the standard deviation of the industry tariff decline is 17.9 percentage points. Industries within manufacturing experienced the largest average tariff cut of 30.2 percentage points, with the average tariff falling from 33.8 to 3.6 percent.

Importantly, these tariff declines significantly affected the volume and structure of Vietnamese exports to the U.S. and worldwide. During this period, Vietnam's aggregate worldwide exports were expanding, but the exports to the U.S. grew even more. Figures 1 and 2, also reported in Fukase (2013), show the value and the share, respectively, of Vietnamese exports to the U.S. from 1997 through 2006. The implementation of the BTA led to a significant surge in exports, which

[16] The BTA required Vietnam to reduce import tariffs on approximately 250 (out of approximately 6000) 6-digit HS agricultural and manufactured food products. As these tariff cuts were small in comparison to the U.S. tariff cuts and only affected a relatively small number of products, we do not discuss them in detail. Our results are robust to controlling for these tariff cuts. As part of the BTA, Vietnam was required to implement various regulatory and legal changes over a period of 10 years following the implementation of the BTA. These included commitments to improve market access in services such as banking and telecommunication, intellectual property rights, and protection of foreign direct investment (STAR-Vietnam (2003)).

[17] McCaig (2011) uses detailed information on U.S. tariffs for both of these tariff schedules from the U.S. International Trade Commission's online Tariff Information Center and computes the ad valorem equivalent of any specific tariffs. He then matches the tariff lines to industries by the concordance provided by the World Bank via the World Integrated Trade Solution database to construct industry-level tariffs according to 2-digit ISIC industry nomenclature. This classification closely matches the industry classification in the VHLSSs.

is evident from the break in trend in 2001 in Figure 1. This break is especially pronounced for manufactured exports, which experienced substantially larger BTA tariff cuts than primary sector exports.[18] Figure 2 indicates that the share of Vietnamese exports going to the U.S. grew rapidly from 5.1 percent in 2000 to 19.0 percent in 2004 and this increase was primarily driven by manufacturing, where U.S. exports accounted for 26.1 percent of Vietnamese exports by 2004.[19] The top eight exports to the U.S. according to 2004 value by industry were apparel; footwear; textiles; food products and beverages; furniture; agriculture; refined petroleum; and office, accounting and computing machinery.

Figure 3 shows the relationship between growth in exports to the U.S. between 2001 and 2004 and tariff changes across 2-digit ISIC industries. A strong negative relationship suggests that industries with greater tariff cuts experienced faster export growth. Appendix Table A.1 reports the industry-level regression of the change in log exports to the U.S. between 2001 and 2004 on the change in U.S. tariffs, which yields a statistically significant estimate of the coefficient on the change in U.S. tariffs for traded industries and for manufacturing. The estimate in column 1 implies that an industry with the average tariff cut, 20.9 percentage points, experienced average annual export growth to the U.S. of 48 percent.

This BTA-related expansion of U.S. exports is not driven by industry-specific global demand shocks. Appendix Table A.1 also reports results for Vietnamese exports to the European Community as an outcome variable. Unlike exports to the U.S., Vietnamese exports to the E.U. were already subject to MFN tariffs prior to the implementation of the BTA (STAR-Vietnam (2003)). As a high-income export market destination, the E.U. likely faces similar industry-specific demand for low-income country exports as the U.S. market. To the extent that U.S. tariff changes are correlated with these shocks, BTA-induced tariff changes would also be spuriously correlated with Vietnamese exports to the E.U. However, the coefficients on tariffs reported in columns 3 and 4 are statistically

[18] Total manufacturing exports also increased following the BTA, as they grew at an annual rate of 23.4% between 2001 and 2006 as compared to 12.8% between 1997 and 2001. The corresponding figures for total exports are 13.1% between 1997 and 2001 and 21.5% between 2001 and 2006.

[19] As a non-member of GATT and the WTO, Vietnam was not subject to the Multi Fibre Agreement and did not initially face any export quotas for textile and apparel products destined for the U.S. In July 2003, a bilateral textile agreement came into force that imposed quotas on Vietnamese textile and apparel exports to the U.S. This agreement is likely responsible for the reduction in the rate of growth of the share of U.S.-bound Vietnamese manufacturing exports following 2003. In the analysis below, this is one of the reasons why we restrict our period to the two years immediately following the implementation of the BTA. To the extent these quotas affected Vietnamese households in 2003 they would likely attenuate our findings.

insignificant and an order of magnitude smaller than the corresponding coefficients in columns 1 and 2, indicating no association between the changes in U.S. tariffs and changes in Vietnamese exports to the E.U. It is therefore unlikely that BTA-induced tariff changes are spuriously correlated with industry-specific global demand shocks for Vietnamese goods.[20]

A fourth useful feature of the U.S. tariff cuts induced by the BTA is that the usual concern about the political economy of protection and the endogeneity of tariff changes are potentially less severe. Industry-specific tariff cuts occurred by the U.S. reassigning Vietnam from one pre-existing tariff schedule to another. Prior to the BTA, imports from Vietnam were covered by Column 2 of the U.S. tariff schedule, whereas after the BTA they were covered by Most Favored Nation tariffs, also known as Normal Trade Relations. The Column 2 and MFN tariffs began to diverge in 1951 when the U.S. assigned Vietnam and twenty other communist countries to a list of countries without normal trade relations. These countries became subject to substantially higher Column 2 tariffs, which were based on tariffs levels legislated by the Tariff Act of 1930 (Pregelj (2005)). The Column 2 tariff rates have remained relatively unchanged over time (Pregelj (2005)). Immediately prior to the BTA, the mean Column 2 tariff across 4-digit HS products remained essentially unchanged, at 31.2 and 31.5 percentage points, respectively in 1997 and 2001, and the correlation was 0.991 (McCaig (2011)). While the U.S. MFN tariffs have fallen over time, Vietnam was not part of the negotiation process as a non-member of the GATT and the WTO.

The U.S. tariff cuts were presented as an all-or-nothing package whereby exports from Vietnam into the U.S. would immediately be covered by MFN tariff rates (negotiated among the WTO members in a round that concluded by 1995) instead of Column 2 tariffs. The movement of Vietnam from one pre-existing U.S. tariff schedule to a second pre-existing U.S. tariff schedule implies that neither U.S. nor Vietnamese industries had an opportunity to influence the tariff cuts faced by specific industries at the time of the implementation of the BTA.

We further confirm the lack of correlation between BTA-induced tariff changes and pre-existing industry trends and levels. In particular, BTA-induced tariff changes do not appear to be related to pre-existing trends in Vietnamese exports to the U.S nor to other high-income

[20] We obtain qualitatively similar results when we exclude industries whose exports accounted for less than 0.5% of total Vietnamese exports in 2001. We also find qualitatively similar results when we use growth rates as in Davis and Haltiwanger (1992) as a dependent variable. These growth rates are defined as $g = (y_t - y_{t-1}) / [0.5(y_t + y_{t-1})]$ and accommodate zero exports in an industry at either the start or end of the period.

destinations such as the E.U. A falsification check of the growth of exports to the U.S. between 1997 and 2000, where the industry-level pre-BTA tariffs are matched with exports in 1997 and the post-BTA tariffs are matched with exports in 2000, yields a coefficient substantially smaller in magnitude that is statistically insignificant (see Appendix Table A.1, Panel B, columns 1 and 2). We obtain a similar finding for growth of exports to the E.U. between 1997 and 2000 (see Appendix Table A.1, Panel B, columns 3 and 4).[21] Thus, the export growth to the U.S. following the BTA is not simply the continuation of pre-existing trends. In addition, we regressed the change in U.S. tariffs on a measure of the unskilled labor intensity of an industry (measured by the share of workers that completed grade 9 or less) and the share of workers within the industry working in household businesses prior to the implementation of the BTA. Across traded, all, and manufacturing industries we find partial correlations of 0.155, -0.120, and 0.030 for the share of unskilled labor and 0.207, 0.047, and 0.056 for the share of informal workers. None of the correlations are statistically significant. Overall, neither contemporaneous growth in demand for Vietnamese exports from other high-income countries, nor pre-existing trends in industry exports, nor baseline industry characteristics are statistically correlated with the BTA-induced industry tariff changes.

4. Data and Aggregate Trends in Household Business Employment

4.1 Definition of a household business

In Vietnam, firms operate either in the household business sector or in the registered enterprise sector. The registered enterprise sector includes four ownership categories: state, collective, foreign, and (domestic) private as defined by the Enterprise Law.[22] All state, collective, and foreign businesses have to legally register as an enterprise. Private businesses can legally operate as either a household business or a registered private enterprise. The legal guidelines for when a private business must register as an enterprise are at times vague, but they consistently require registration as an enterprise for private businesses that regularly employ workers, or employ more than 10 workers, or that operate in more than one location.[23] Thus, while small, single-location businesses may operate as household businesses or enterprises, all larger businesses

[21] A similar regression for worldwide exports between 1997 and 2000 also yields statistically insignificant findings.
[22] See law No. 13-1999-QH10 Law on Enterprises.
[23] Decrees No. 02/2000/ND-CP of 3 February 2000 and No. 109/2004/ND-CP of 2 April describe household business and enterprise registration requirements during our study period, with the first decree focusing on regular employment and the second on the 10 worker threshold.

are required to operate as enterprises. Note that the average household business in manufacturing has only 1.5 workers (including the owner), well below the enterprise employment threshold, and being a household business does not imply that a business operates illegally. Household businesses can operate in the physical premise of a household (or farm), market stalls, industrial zones, trade centers, and in variable locations (e.g., street vendors).

While the definition of an informal business varies across countries, using the distinction between a household business and an enterprise in Vietnam is consistent with informal firm definitions in other countries. Commonly, the informal versus formal distinction is about firm registration status, which may be related to other firm characteristics.[24] La Porta and Shleifer (2008) define informal as unregistered with the central government and thus without a tax identification number in a collection of informal and micro firm surveys across countries. Although the exact legal requirements and number of steps to formally register vary across countries, firm informality is regularly based on the firm's registration status for a license, certificate, or tax code.[25] Informal firms frequently do not hire workers and a large component of the informal firm distribution is self-employment and own-account work.[26] In Section 5.2, we therefore explore the link between the concept of household business and self-employment and check the robustness of our findings to the use of self-employment.

In Vietnam, most household businesses are household farms in agriculture and aquaculture. Non-farm household businesses predominately operate in services (70%). Of the 30% of household businesses in manufacturing, the most common activities are production of food and beverages, wood processing, clothing, furniture, and textiles. The difference in registration status is predictive of important differences in underlying firm characteristics in the household business and enterprise sectors. Consistent with other studies on household businesses (La Porta and Shleifer (2008, 2014) and Nataraj (2011)), Vietnamese household businesses are substantially smaller and have lower

[24] For example, manufacturing firms in India that have 10 or more workers and use electricity or have 20 or more workers but don't use electricity are required to register (Nataraj (2011)).

[25] A related literature, primarily from Latin American countries and studying domestic import liberalization, focuses on whether formal firms comply with labor legislation requirements concerning contracts, severance payments and social insurance contributions for workers. Some of these studies face data constraints. Goldberg and Pavcnik (2003) and Bosch, Goñi-Pacchioni, and Maloney (2012) use survey data covering only urban areas; Paz (2014) uses a nationally representative household survey, but excludes the self-employed from the analysis; and Menezes-Filho and Muendler (2011) use matched employer-employee data that covers workers with a signed work card in registered firms and the same urban survey as Goldberg and Pavcnik (2003).

[26] Gollin (2008), Banerjee and Duflo (2007), de Mel, McKenzie, and Woodruff (2009), Fafchamps, McKenzie, Quinn, and Woodruff (2014) equate informal microenterprises with self-employment or own-account work.

labor productivity than firms in the enterprise sector. For example, the average household business in manufacturing has only 1.5 workers (including the owner), while the average employment size for manufacturing firms in the enterprise sector is 152.[27] Household businesses in manufacturing have on average 6 times lower average revenue per equivalent hour worked than enterprises (see Section 6 and Appendix B for further details on this calculation).

Registered enterprises are required by the Enterprise Law to follow formal accounting standards and to report comprehensive information about their financial position, including information on their workforce. Consequently, as in other low-income countries, in Vietnam workers in the enterprise sector are captured in the conventional firm-level datasets covering the formal sector, whereas workers in the household business sector are not. The next section describes how we use comprehensive household surveys to observe workers in both sectors.

4.2 Data Description

We use two waves of the Vietnam Household Living Standards Surveys (VHLSS) conducted by the General Statistics Office (GSO) of Vietnam in 2002 and 2004 as our primary data source. The surveys are nationally representative, have a 12 month recall, and cover 2001/2002 and 2003/2004.[28] While the VHLSS is a repeated cross section of households, it also contains a smaller panel subsample, which we employ in several specification checks.[29]

We focus on employed individuals, ages 20 to 64, in their main job (i.e., the most time consuming). [30,31] We create variables on workers' demographic and educational characteristics

[27] These estimates are based on the estimated number of manufacturing household businesses and primary job workers in these businesses from the 2002 VHLSS for household businesses and the end of year employment in the 2001 Enterprise Survey for manufacturing enterprises.

[28] The BTA was implemented on December 10, 2001. The 2002 survey interviewed households throughout the year. With a recall period of 12 months, individuals interviewed at the start of 2002 have a recall period that almost entirely precedes the BTA, while individuals interviewed at the end of 2002 have a recall period almost exclusively after the implementation of the BTA. Our results thus potentially underestimate the full impact that the BTA has had on labor reallocation as of 2003/04.

[29] In robustness checks, we rely on two additional data sources: the 1992/93 and 1997/98 Vietnam Living Standard Surveys (VLSS), predecessors to the VHLSS. Additionally, we use Vietnam's Enterprise Survey for 2000 and 2003, a firm-level dataset that covers all registered firms in the enterprise sector, for an analysis of industry employment changes.

[30] For each individual in the household the survey collects information on whether the individual is employed, unable to find work or out of the labor force. Unemployment is very infrequent in our data. For example, among individuals age 20 to 64 in the 2004 VHLSS, 88.7 percent report working during the past 12 months while only 6.2 percent of those not working (or 0.7 percent of the age group) report being unable to find a job.

(gender, age, highest level of completed education, ethnic minority status), geographic location (urban residence, province), and industry affiliation. The survey distinguishes between 60 2-digit ISIC (Rev 3) industries overall, 34 in the traded sector, and 22 of which are in manufacturing. We use industry affiliation to link individual-level data to industry-level U.S. tariffs on Vietnamese exports, described in Section 3.

We construct the main variable of interest, an indicator for whether a worker works for a household business, from a survey question about the worker's employer type. The question distinguishes whether a worker is self-employed, works for another household, the state sector, the collective sector, the private enterprise sector, or the foreign sector. The indicator takes the value one if an individual works in his/her own household business or in another household's business, and zero otherwise. This definition of employment in a household business is consistent with the distinction between household businesses and registered enterprises as per Vietnam's Enterprise Law as discussed in Section 4.1.[32]

One potential problem with the construction of a household business indicator is that the individuals might not know whether they work for a household business or a private enterprise. While this is a concern, the survey provides detailed instructions to the enumerators about how to record the answers to questions. Furthermore, most workers in household businesses work for their own business and presumably know its registration status.[33] If measurement error was severe, one would not expect to observe differences in worker outcomes such as earnings and benefits for workers in household businesses and other establishments. As we discuss below, we find notable differences in wages and benefits received between workers in the household business and enterprise sectors.[34] To the extent that there is some measurement error in our dependent variable, it would reduce the precision of our estimates and bias us toward finding no significant

[31] Among workers age 20 to 64 in the 2004 VHLSS, 42.5 percent reported working more than one job during the past 12 months. Among these individuals the average annual hours worked was 1355 and 511 in their primary and secondary jobs respectively as compared to 1907 hours for workers that reported working only one job.
[32] The 2004 VHLSS distinguishes between self-employment in a household business and self-employment in a private enterprise, while the 2002 VHLSS does not. To be consistent across surveys we classify all self-employed individuals as working for a household business. This is not a very egregious grouping since self-employment in the private sector is only 0.7 percent of self-employment across all industries and 1.6 percent of self-employment in manufacturing in the 2004 VHLSS.
[33] In section 5.2, we also rely on an indicator for whether an individual is self-employed.
[34] This also holds if we compare workers in household businesses and private enterprises.

impact. Appendix Table A.2 provides summary statistics for the sample of 152,388 workers in 2001/02 and 96,407 workers in 2003/04.

4.3 Employment in enterprises versus household businesses

While our study can capture worker allocation between employers in the household business and enterprise sectors – a margin that is not observed in firm-level or matched employee-employer administrative data – we do not observe the allocation of workers across firms within employer types. Our study thus complements the literature on labor allocation across heterogeneous employers in the formal sector.

Household businesses tend to be substantially less productive than firms in the enterprise sector, even relative to smaller private enterprises, a point we return to in Section 6. Large labor productivity gaps, combined with a large employment share of informal firms in many low-income countries, suggest that focusing on this margin of labor adjustment may be important for aggregate productivity.

However, the distinction between a household business and an enterprise is also important from the perspective of the workers. Appendix Table A.3 reports summary statistics on several worker-related outcomes for individuals employed in household businesses and enterprises. To begin with, self-employment is very high in the household business sector (83 percent of household business workers in all industries and 61 percent in manufacturing industries are self-employed). For wage earners, hourly wages are higher in the enterprise sector. These wage gaps persist when one compares observationally equivalent workers.[35] For example, manufacturing workers in a household business earn about 25 percent less per hour than observationally equivalent workers working in the same industry and province (Column 6 of Appendix Table B.1). Controlling for unobserved worker characteristics, informal manufacturing workers that switch to work for an enterprise tend to earn 9 percent more than when they work for a household business (Columns 1 and 2 of Appendix Table B.2). These patterns for Vietnam are consistent with the literature on firm size and earnings and on informality (Marcouiller, Ruiz de Castilla, and Woodruff (1997) and Goldberg and Pavcnik (2003)). Similar differences emerge in earnings, which include income from self-employment (see Appendix Table B.3).

[35] See Appendix B.1.3 for detailed discussion.

Workers that work for household businesses are less likely to report receiving non-wage/salary payments, such as for holidays (see Appendix Table A.3). Additionally, workers in the enterprise sector are legally entitled to compulsory social insurance contributions on their behalf by their employers. This is in addition to their contemporary wage/salary payments. Hence, these workers will have access to a formal pension upon retirement. Unfortunately, our data does not include information on whether an employer provides these benefits, so we cannot measure such compliance.

Finally, Appendix Table A.3 shows that enterprise workers work longer hours in their primary job (about 25% more per year across all industries) and are substantially less likely to work more than one job. This suggests that precarious work is less of a concern for these workers. Furthermore, Banerjee and Duflo (2008, 2011) discuss psychological benefits of secure employment. An important difference between being employed in the formal versus the informal sector is that workers in the formal sector have more stable jobs from the perspective of hours worked in a given week (as opposed to having to piece together hours across one or two jobs). Overall, this discussion suggests that the distinction between employment in the household business and enterprise sectors has additional implications for workers than simply the difference between working for a more and/or less productive firm in the formal sector would.

4.4 Aggregate trends in household business employment

Table 2, Panel A reports the aggregate share of individuals that work in household businesses in Vietnam in 2001/02 and 2003/04 and motivates the importance of this employment margin.[36] The results are presented for workers in all industries, in industries other than agriculture and aquaculture, and in manufacturing. The major fact to emerge is that employment in household businesses is very high in Vietnam. Economy-wide, 85 percent of workers are employed in household businesses in 2001/02. The prevalence of employment in household businesses does not merely reflect the large overall share of employment in agriculture and aquaculture, as the share continues to be high, at 67 percent, when we exclude agriculture and aquaculture.[37] We observe similarly high levels of working for household businesses, 66 percent, within manufacturing,

[36] See McCaig and Pavcnik (2015) for descriptive analysis of the decline in the share of workers in informal firms and patterns of birth cohort and individual attachment to the informal sector over a decade.
[37] The middle panel also excludes forestry, a very small sector; for brevity, we refer to agriculture and aquaculture only.

consistent with evidence from India (Nataraj (2011)) and Ghana (Gollin (2008)). Thus, even in manufacturing, the sector that is the focus of most of the existing work on trade and labor allocation, the usual analysis of formal enterprise firms captures a small share of employment.

The second key fact to emerge from Table 2 is the decline in the prevalence of working in household businesses between 2001/02 and 2003/04. Economy-wide, the share of workers in household businesses fell by 3.3 percentage points (or 4 percent). The drop was particularly pronounced in manufacturing, where the share of workers employed in household businesses fell by 5.6 percentage points (or 9 percent). The conceptual framework in Section 2 emphasizes that trade can influence the composition of employment through the reallocation of employment across employers within industries and between industries with differential prevalence of household business employment. We examine whether the observed aggregate changes in the incidence of employment in household businesses stem from changes in the structure of employment across industries (e.g., expansion of employment in industries that tend to organize their production in formal enterprises) or from within-industry reallocation of workers across employers. We decompose the change in the share of workers in household businesses in total employment between 2001/02 and 2003/04, denoted by ΔH , into within and between industry shifts, respectively: $\Delta H_t = H_t - H_{t-1} = \sum_j \Delta h_{jt} s_j + \sum_j \Delta s_{jt} h_j$, where s_{jt} is the share of industry *j*'s employment in total employment at time *t*, h_{jt} is the share of workers in household businesses in total employment in industry *j*, $s_j = 0.5\left(s_{jt} + s_{jt-1}\right)$, and $h_j = 0.5\left(h_{jt} - h_{jt-1}\right)$. The first summation term captures the importance of mobility of workers across employers *within* an industry and the second summation term captures the prevalence of mobility of workers *across* industries as sources of changes in aggregate employment in household businesses.

Panel B of Table 2 presents the results of the decomposition. Economy-wide, both channels contribute equally toward the decline in the aggregate share of household business employment. The between-industry component accounts for 48 percent of the aggregate decline and mainly reflects the relative contraction of employment in agriculture and aquaculture, where almost all workers work in household farms. Exclusion of agriculture and aquaculture raises the contribution of the within-industry channel from 52 to 86 percent. The within-industry reallocation of workers across employers from the household business to the registered enterprise sector plays an even larger role

in manufacturing, where it accounts for the entire decline in the aggregate share of household business employment. Overall, these aggregate trends motivate our empirical analysis, which we turn to next.

5. Empirical Implementation

This section first describes our empirical methodology and main results, followed by discussing several robustness and falsification checks. The section concludes with a discussion of the implications of the BTA for industry employment.

5.1 Empirical Methodology and Main Results

We exploit large heterogeneity across industries in declines in U.S. tariffs on Vietnamese exports induced by the BTA to investigate the relationship between exporting costs and the allocation of workers between employers in the household business and enterprise sectors. The empirical methodology relies on a comparison of the probability that a worker works for a household business before and after implementation of the BTA across Vietnamese industries differentially exposed to the declines in U.S. tariffs. In the initial empirical specifications, we estimate the following linear probability model:

$$H_{ijt} = X_{ijt}\delta + \beta tariff_{jt} + \gamma_p + \lambda_j + \theta_t + \varepsilon_{ijt}. \quad (1)$$

H_{ijt} is an indicator for whether a worker *i* employed in industry *j* at time *t* works for a household business, X_{ijt} is a vector of worker characteristics (this vector includes age, age squared, and indicators for highest completed education category (primary, lower secondary, upper secondary, with no formal education as the excluded category), gender, ethnic minority indicator, an indicator for whether a person lives in a rural area), $tariff_{jt}$ is the U.S. tariff on Vietnamese exports in industry j at time t. The specification also includes province (γ_p), industry (λ_j), and time (θ_t) fixed effects. The main parameter of interest is the coefficient on tariffs. A positive coefficient implies that a decline in tariffs is associated with a decline in the probability of working in a household business and the reallocation of labor away from household businesses. Standard errors are clustered by industry to account for general forms of heteroskedasticity and serial correlation in the error term within an industry.

Inclusion of individual worker demographic characteristics in equation (1) sweeps out differences in worker composition across industries, employers, and time that could simultaneously affect the allocation of labor and be spuriously correlated with tariff levels. The post-BTA fixed effect controls for aggregate economy-wide adjustments in household business employment coinciding with the implementation of the BTA agreement. Province fixed effects absorb any time-invariant features of provinces affecting labor market conditions in a province, while industry-level fixed effects capture all time-invariant industry characteristics correlated with tariff levels and prevalence of household business employment. In this set up, the empirical strategy identifies the coefficient on tariffs by comparing effects of tariff declines on workers with the same observable characteristics within provincial labor markets, some of whom worked in industries that experience large tariff cuts and others who worked in industries with smaller tariff cuts.

Any potential threats to the underlying identification assumption would stem from industry specific time-varying factors that covary with industry tariff changes and independently influence industry-specific changes in the propensity to work for a household business. As discussed in detail in Section 3, the institutional implementation of the BTA-induced tariff cuts eliminated the ability of industry-specific contemporaneous conditions in Vietnam or the U.S. to influence the magnitude of industry tariff cuts through the political economy of tariff formation. One could potentially still be concerned about spurious correlation between industry tariff changes and contemporaneous industry-specific changes in global demand for Vietnamese exports. In Section 3, we also show that U.S. tariff changes are not spuriously correlated with contemporaneous industry-specific changes in global demand for Vietnamese exports: the U.S. tariff declines lead to a strong increase in Vietnam's exports to the U.S, but are not associated with changes in export growth to the E.U. This also likely eliminates the role of contemporaneous supply shocks in Vietnam, which would be affecting all global destinations. We also find no statistically significant association between U.S. tariff changes and industry baseline characteristics, such as the share of household business workers in industry employment and the unskilled-labor intensity of the industry, prior to the implementation of the BTA, nor between U.S. tariff changes and pre-existing industry-specific time trends in Vietnamese exports to the U.S., the E.U., and worldwide. These results, discussed in detail in Section 3, further validate the identification strategy in equation (1). Finally, we focus on the short run effects of the BTA because we want our identification strategy to only capture the effects of the BTA as opposed to other changes occurring in the Vietnamese economy, including World

Trade Organization (WTO) accession in 2007 and the 2008 crisis. Our estimates should therefore not be viewed as long-run estimates of the effects of the BTA, which are potentially larger (see Dix-Carneiro and Kovak (2017) for long-run adjustment to import liberalization).

Figures 4 and 5 present scatterplots of the change in the share of household business workers in an industry and the BTA-induced change in U.S. tariffs for all traded industries and for manufacturing, respectively.[38] The size of the circles reflects the employment size of each industry. The slope of the displayed regression lines is equivalent to the estimate of the coefficient on tariffs based on equation (1) without controlling for worker characteristics and province fixed effects.[39] The figures show a clear positive relationship: industries with larger tariff cuts experienced larger reductions in the share of workers working in household businesses.

The relationships shown in the scatterplots continue to hold once we estimate the coefficient on tariffs as specified in equation (1) and reported in Table 3. Column 1 presents estimates of equation (1) for traded industries. We find that workers in industries that faced greater reductions in U.S. tariffs experienced larger decreases in the probability of employment in household businesses relative to observationally equivalent workers in industries with smaller tariff reductions. The magnitude of the coefficient (0.209) suggests that an industry that experienced the average reduction in tariffs, 20.9 percentage points, saw the probability of working in a household business fall by 4.4 percentage points relative to an industry facing no reduction in tariffs. In Column 2, we report the estimates of equation (1) for workers in all industries, including non-traded industries. The non-traded sectors were not directly impacted by the tariff cuts and observed no change in tariffs.[40] The inclusion of non-traded sectors dampens the magnitude of the coefficient relative to the estimate based on the traded sector alone, although the coefficient continues to be positive and statistically significant. Lastly, in column 3 we estimate equation (1) for the manufacturing sector, a sample that is more comparable to the samples used in most studies of labor reallocation in response to trade reform. The estimated coefficient suggests that the average reduction in manufacturing tariffs of 30.2 percentage points is associated with a 4.7 percentage

[38] Both figures exclude industry 12 (mining of uranium and thorium ores) from the display, but not from the regression line, as it is an extreme outlier and a very small industry in terms of employment.

[39] The industry observations are weighted by $n_j^{2002} n_j^{2004} / (n_j^{2002} + n_j^{2004})$ where n_j is the number of workers in industry j in the indicated year.

[40] We assign a tariff of 0 to non-traded industries in both years. Equation (1) includes industry fixed effects, which implies that non-traded industries experience no tariff change. See Kovak (2013) for an alternative approach in the local labor markets literature.

point reduction in the probability of employment in a household business in that industry.[41,42] Importantly, our estimates reflect short run responses and may underestimate the long run effects of the BTA on employment in the formal sector, as the economy has more time to adjust.

The results in Panel A of Table 3 are robust to a variety of specification checks. In Panel B we report estimates based on a specification that replaces the province and year fixed effects with province-year fixed effects. In addition, in Panel C we also allow the effects of worker observables to vary over time by interacting the individual covariates with a 2004 indicator. The estimates in Panel B and C are very similar to those in Panel A.[43] In Appendix Table A.4, we report regression results from additional modifications of equation (1). The specification in Panel A removes all individual covariates from equation (1). The specification in Panel B estimates equation (1) with a sample that excludes observations from mining of uranium and thorium ores (industry 12), a small industry, but a significant outlier. The results remain consistent across these additional specifications.[44]

The analysis thus far focuses on the extensive margin of labor adjustment. As workers move into the enterprise sector they may not work the same number of hours as current enterprise sector workers, so that labor reallocation measured in terms of hours may differ relative to reallocation solely on the number of workers. We compute the share of total hours worked in the household business sector in an industry using information on hours worked in the primary job. In the spirit of Figures 4 and 5, we use this variable as a dependent variable and estimate an industry-level version of equation (1) (without individual covariates and province fixed effects), weighting each industry by its average size. The results are reported in Appendix Table A.5. Consistent with

[41] The magnitude is slightly smaller for manufacturing than for all traded industries because of agriculture. In column (1), the coefficient on tariffs is identified by differential changes in household business employment across industries, including agriculture, which received a lower tariff reduction than most manufacturing industries. During this period, households, not enterprises, undertake essentially all agricultural activity in Vietnam. Consequently, regressions in column 1 that include agriculture are estimated with additional observations that, relative to observations from manufacturing, tend to experience almost no change in the share of household business employment and a small tariff decline. This contributes to higher magnitude of coefficient in column 1 than 3.

[42] These results are robust to controlling for Vietnam's BTA tariff reduction commitments, which are concentrated in agriculture and the processing of food and beverages. The estimated coefficient on U.S. tariffs is 0.170, 0.131, and 0.180 on traded, all, and manufacturing industries respectively, all of which remain statistically significant at the 1 percent level.

[43] To the extent that different price changes of non-tradables across provinces (as in Kovak (2013)) would be key for our results in column 2, the estimates of the effects of the BTA in column 2 would change substantially in Panel B of Table 3, relative to Panel A. The estimated coefficients are similar.

[44] Our main findings are also robust to estimation using probit or logit. Results are available upon request.

our main findings, the results document a reduction in the share of hours worked in the household business sector in response to the tariff cuts. Moreover, the similar magnitudes of the coefficients in Appendix Table A.5 and Panel A of Table 3 suggest that as workers move between sectors, the average number of hours worked per person within each sector is relatively unchanged. Hence, the primary margin of adjustment in hours worked within an industry appears to be reallocation of workers across sectors.

5.2 Results and Falsification Test based on Self-Employment

One may worry that our results might not generalize to other settings because our definition of employment in a household business is specific to the definition of firm informality in Vietnam. The discussion in Section 4.1 illustrates that this definition is consistent with those for many other countries studied in the literature on informal microenterprises. In addition, this definition is highly correlated with self-employment (in these microenterprises) in low-income countries. For example, La Porta and Shleifer (2008) show that self-employment correlates highly with various measures of informality in a large set of low-income countries. This is also the case in Vietnam, where the correlation between self-employment and working in a household business is 0.67.

Importantly, our main finding that declines in tariffs on exports are associated with reductions in informality (i.e., declines in probability of working for a household businesses) is robust to using self-employment as a dependent variable. We estimate equation (1) with an indicator for self-employment as the dependent variable and report the estimated coefficients on tariffs in columns 1 to 3 in panel A of Table 4. The coefficients are positive, statistically significant, and of similar magnitudes as the corresponding coefficients on tariffs in Table 3. The similar magnitudes of the coefficients suggests that movement out of working for household businesses reflects both movements from self-employment and wage work in a household business, although movements out of self-employment play a slightly stronger role in manufacturing. The margin of self-employment versus paid employment is potentially more comparable across countries than definitions of informality based on the country-specific legal definition of an informal firm and this margin is more commonly available in labor force or household surveys in low-income countries.

We further focus on self-employment to show that the previous results are not driven by differential pre-existing employment trends across industries that differ in their propensity to

organize production in household businesses. As discussed in Section 3, the industry changes in U.S. tariffs are not related to initial industry conditions, such as the share of household business workers within an industry or industry skill intensity, nor to pre-BTA growth in exports to the U.S. A falsification test that uses two rounds of data covering a pre-reform period further finds no evidence that changes in industry tariffs are correlated with pre-existing trends in household business employment across industries. We perform this test using information from the 1993 and 1998 Vietnam Living Standards Surveys (VLSSs) and assign the pre-BTA tariffs (Column 2 tariffs in 2001) to the 1993 data and the post-BTA tariffs (MFN tariffs in 2001) to the 1998 data.[45] The employment module for the 1993 VLSS does not separately identify employment in a household business from employment in a private sector business because Vietnam did not make a legal distinction between household businesses and private enterprises at that time. Consequently, we use an indicator for being self-employed as the dependent variable.

If pre-existing trends in household business employment were correlated with industry-specific U.S. tariff cuts, this specification would yield estimates of tariff coefficients of the same sign and similar magnitude to the coefficients obtained in the corresponding analysis using data surrounding the actual policy change. The results are presented in panel B of Table 4. The estimated coefficients on tariffs are close to zero in magnitude, always statistically insignificant, and differ from the estimates of the corresponding coefficients in panel A of Table 4 based on data surrounding the period when BTA was actually implemented.[46] Underlying trends therefore cannot account for the strong relationship between the U.S. tariff reductions and the decrease in the probability of working for a household business that we reported in Table 3 and the top panel of Table 4, further validating the identification strategy.

We further examine the robustness of our findings to industry-specific pre-existing trends by including these trends directly in our main specification, equation (1), which uses an indicator for

[45] The 1993 and 1998 VLSSs are based on the same sampling framework, which differs from the sampling framework used for the 2002 and 2004 VHLSSs.

[46] The industry codes between the 1993 and 1998 VLSSs do not perfectly match. In particular, 2-digit ISIC revision 3 industries 31 and 32, 34 and 35, and 30 and 33 were merged together since the 1993 VLSS used a more aggregate industry definition in these instances. Additionally, industries 17 and 18 and 20 and 36 have also been merged since the 1998 VLSS appears to have switched the assignment of some workers in some of these industries. The low point estimates of the tariff coefficient and the lack of statistical relationship in panel B of Table 4 (relative to panel A) do not simply reflect higher levels of industry aggregation. When we estimate the specifications in panel A at the same level of industry aggregation as the bottom panel, we continue to obtain positive and statistically significant coefficients on tariffs during the period that spans the BTA (0.174 (0.015) for traded in column 1, 0.101 (0.036) for all industries in column 2, and 0.159 (0.018) for manufacturing in column 3).

working in a household business as a dependent variable. In particular, we add three pre-BTA industry-specific trends: the change in ln employment, the change in the self-employment rate, and the change in the mean grade completed, all computed between 1993 and 1998, interacted with a 2004 indicator as controls to the specification in (1). We report the results in Panel C of Appendix Table A.4. The table reports the coefficients on tariffs, as well as the coefficients on the included pre-existing trends interacted with the 2004 indicator. The tariff coefficients are similar to those we report in Table 3, especially for traded and manufacturing industries.

5.3 Heterogeneity in Worker Responses to Tariff Declines

The results show that large BTA-induced declines in industry-specific export costs decrease the probability that Vietnamese workers work for a household business, leading to a reallocation of workers toward the formal enterprise sector. The overall effects analyzed so far might mask heterogeneity in responses of workers. We explore this potential heterogeneity to tariff cuts by location, age, gender, and education.

Vietnamese provinces differ in their degree of integration with international markets and this heterogeneity in part reflects proximity to a major seaport. Provinces closer to major seaports are more internationally integrated and more exposed to export opportunities (World Bank, 2011). For example, the information on the value of manufacturing exports from the 2000 Enterprise Survey suggests that five provinces with or near major seaports (Ho Chi Minh City, Dong Nai, Hanoi, Binh Duong, and Hai Phong) account for over three quarters of reported manufacturing exports. To the extent that export opportunities associated with the BTA disproportionately increase labor demand in the larger firms operating in the export sector, as noted in the conceptual framework in Section 2, one would expect a relatively larger increase in labor demand among firms in the enterprise sector in more integrated provinces. Consistent with this view, McCaig (2011) finds that poverty declined and average wages increased relatively more in provinces with a higher concentration of export-oriented industries at the onset of trade reform. The impact of U.S. tariff cuts on the incidence of household business employment would then be expected to be more pronounced in more internationally integrated provinces.

To explore the possible heterogeneity of effects by location, we split Vietnam's provinces into two groups based on the median distance from one of Vietnam's three major seaports in Hai Phong, Da Nang, and Ho Chi Minh City. We estimate equation (1) for each sample. The results are

presented in Panel A of Table 5. As expected, declines in U.S. tariffs are associated with larger relative declines in household business employment for individuals living in more internationally integrated provinces. While all estimates of the coefficient on tariffs are positive, the magnitudes of the coefficients are substantially larger and always statistically significant in provinces closer to major seaports. The difference in magnitude and statistical significance of the estimated coefficients on tariffs is particularly notable in manufacturing, the sector most exposed to the BTA tariff cuts. This finding is consistent with Hanson (1996).

We also examine heterogeneity in responses to tariff cuts by worker age, gender, and education. This heterogeneity could stem from differences in adjustment costs across workers with different demographic characteristics (see Dix-Carneiro (2014), Coşar (2013)), or it could reflect differential changes in labor demand across worker types. These results are also presented in Table 5. We split workers into five age groups and estimate equation (1) separately for each of the groups. The probability of working in a household business declines more for young workers in response to the US tariff cuts (column 1) in the traded sector and economy-wide (column 2). The heterogeneity in responses to tariffs by age appears at first less pronounced in manufacturing (column 3). However, the implied share of reallocated young workers is above, while the implied share of reallocated workers in older age groups is below the predicted share of reallocated workers manufacturing-wide. Gender does not appear to differentially affect the responsiveness of working in a household business to tariff cuts. Estimates of equation (1) by gender in Table 5 suggest that men and women were similarly affected by tariff declines. We also estimate equation (1) separately for three education groups: 0 to 8 years of formal education (i.e., did not complete lower secondary), 9 to 11 years of formal education (i.e., completed lower secondary, but not upper secondary), and 12 or more years of formal education (i.e., completed upper secondary). We consistently find that workers with a medium level of education observed smaller declines in the probability of working for a household business than workers with low or high levels of education, albeit these differences are not statistically distinguishable. The larger response of workers with the highest level of education is consistent with models and empirical evidence of increased demand for highly educated workers among exporting, formal firms, particularly when exporting to high-income countries (Bustos (2011a, b), Verhoogen (2008), Brambilla, Lederman, and Porto (2012)) and potentially lower adjustment costs of the more educated (Dix-Carnerio (2014)).

Note that workers that live in more internationally integrated provinces, younger workers, and more educated workers are less likely to work in the household business sector even prior to the BTA. One implication is that the declines in export costs further enlarge the gap in the probability of working in a household business between workers in provinces that differ in access to international seaports and between younger and older workers. This gap also widens between the middle and the highest education group, but narrows between the low and middle education group.

5.4 Longitudinal analysis

The VHLSS resurveyed about 30 percent of the households from 2001/02 in 2003/04.[47] Using this smaller longitudinal subsample, we examine the robustness of the results to selection on unobserved individual-level heterogeneity into moving out of household businesses. We restrict the analysis to individuals aged 20-64 in 2001/02 that worked in both years.[48] We estimate a version of equation (1):

$$H_{ijt} = \alpha_i + \beta tariff_{jt} + \lambda_j + \theta_t + \nu_{ijt} \qquad (2)$$

where the vector of individual characteristics and province fixed effects have been replaced by an individual fixed effect α_i.

To establish comparability with the results from Section 5.1, we first estimate the specification in equation (1) using the longitudinal subsample. The results are reported in columns 1 to 3 in panel A of Table 6 and confirm the findings from Table 3. With the exception of manufacturing, the magnitudes of the coefficients based on the longitudinal sample are somewhat lower than the magnitudes of the corresponding coefficients based on repeated cross-sections, but they are not statistically different from each other.[49] The specifications so far use the tariff in a

[47] The household panel is based on a random selection of enumeration areas from the 2002 VHLSS.

[48] In order to be part of the household panel, the household, or at least some of its members, must reside in the same location as in 2001/02. 9.7 percent of individuals in panel households that report working in 2001/02 are not in the sample in 2003/04. The attrited individuals are more likely young and better educated, and were more likely initially employed in the enterprise sector, in industries that received larger tariff cuts, and by an enterprise in industries that received larger tariff cuts. The attrition might thus bias the coefficient on tariff in the panel estimation downward toward zero, increasing the likelihood of finding no relationship.

[49] Lower magnitudes of the coefficients based on longitudinal subsample could reflect attrition and slight differences in the composition of the longitudinal and cross-sectional samples. As discussed earlier, attrition could potentially bias our estimates downward. Second, to track the same individuals over time, the longitudinal sample

worker's contemporaneous industry at time *t* as a measure of exposure to industry export costs. In the longitudinal data, workers' exposure to export costs can also be measured based on the workers' initial industry of employment, further allowing one to control for the sorting of individuals across industries. Panel B of Table 6 reports estimates of equation (1) based on the tariffs in the worker's initial industry of employment. The magnitudes of the coefficient on tariffs are similar to those obtained in panel A with the contemporaneous industry tariff. In the remainder of this section, we measure workers' exposure to the export cost shock through the initial industry of employment.

The empirical strategy so far compared effects of tariff declines on workers with the same observable characteristics within provincial labor markets, some of whom worked in industries that experienced large tariff cuts and others who worked in industries with smaller tariff cuts. Note that to the extent that workers might select to work in the enterprise sector because of higher expected earnings in this sector based on observable characteristics such as education, gender, age, and minority status included in specification in equation (1), we already account for selection through the inclusion of direct controls for such observable worker characteristics. In addition, any form of selection is only a concern to the extent that it is industry-specific and spuriously correlated with BTA-induced industry-specific tariff changes. The specification in equation (2) includes worker fixed effects, directly controlling for time-invariant individual-level heterogeneity in unobserved worker characteristics that might influence the selection of workers into industries and the propensity to switch employers. The estimates from this specification are reported in columns 1-3 in panel C of Table 6 and confirm the existing findings. Individuals initially working in industries that experience larger tariff cuts face greater declines in the probability of working for a household business than observationally equivalent individuals initially working in industries with lower tariff cuts. The inclusion of individual fixed effects somewhat reduces the estimate of the coefficient on tariffs. For example, the magnitude of the coefficient on tariffs for traded sectors falls from 0.146 to 0.112, implying that a 20.9 percentage point decline in tariffs was associated with a 2.4 percentage point decline in the probability of household business employment. In manufacturing, the coefficient on tariffs drops from 0.179 to 0.09, so that a 30.2 percentage point decline in tariffs is associated with a 2.7 percentage point decline in the probability of working for a household business. Overall,

includes individuals based on initial age (ages 20-64 in 2002), not contemporary age and excludes individuals that enter or exit the workforce because we only have one observation for their work status.

greater declines in exporting costs are associated with greater reallocation of workers from household businesses to employers in the enterprise sector, although the magnitudes of the effects are attenuated in manufacturing.

5.5. Implications for Industry Employment

The literature on the effects of trade on net industry employment finds limited or no response to import tariff declines in the short run in less developed countries (see Goldberg and Pavcnik (2007) for a survey). These studies primarily focus on the consequences of domestic trade liberalizations, which reduced tariffs on imported goods entering the domestic market. Data constraints precluded many earlier studies from examining the effects of trade policy on industry employment patterns representative of the nationwide labor force.[50] Our data provides comprehensive coverage of workers in all industries, in both formal and informal firms, and in urban and rural areas. In this section, we reexamine the effect of industry trade costs on the structure of total employment across industries with this comprehensive dataset and in a setting where trade liberalization primarily lowered tariffs on exports in a destination market

We relate industry tariffs to the structure of employment across industries by estimating the following specification:

$$s_{jt} = \beta tariff_{jt} + \lambda_j + \theta_t + u_{jt} \quad (3)$$

where s_{jt} is the share of industry j at time t in total employment and all other notation follows previously introduced notation. The results, based on estimating equation (3) with an industry's employment share in total employment as a dependent variable (i.e., employment in household businesses and enterprises), are presented in Panel A of Table 7. Interestingly, the magnitude of the coefficients on tariffs is virtually zero and always statistically insignificant when we consider changes in the overall industry structure of employment. These findings are consistent with those of Feliciano (2001) for Mexico and Attanasio, Goldberg, and Pavcnik (2004) for Colombia, which use household survey data that are representative of the entire urban workforce and find no evidence of changes in total industry employment in response to changes in trade policy. Those studies use household surveys that only cover urban areas. For comparison purposes, we replicate the analysis

[50] Feliciano (2001) and Attanasio, Goldberg, and Pavcnik (2004) rely on household surveys that are representative of urban areas, while Revenga (1997), Currie and Harrison (1999), Wacziarg and Wallack (2004), and Topalova (2010) use data covering employment in formal manufacturing firms above a certain employment cut-off.

from Panel A of Table 7 using only urban households. These results are reported in Appendix Table A.6 and yield similar results to using all households. Thus, our findings on the effects of trade policy on the structure of total employment across industries are consistent with the findings from existing literature.

Moreover, our analysis in Sections 5.1 through 5.4 highlights compositional changes across household business and enterprise employers within industries. Further analysis shows that the structure of employment across industries is shifting toward industries with larger tariff cuts in the enterprise sector, the sector most directly impacted by export liberalization. In particular, we estimate equation (3) with industry employment shares obtained from the Enterprise Survey, which covers all firms in the enterprise sector.[51] The results are presented in Panel B of Table 7. The negative estimates of the coefficients on tariffs suggest greater expansion in enterprise sector employment in industries with larger tariff cuts. Importantly, the coefficients in Panel B are at least an order of magnitude larger than the corresponding coefficients obtained for the overall industry employment in Panel A.[52] Thus, the structure of industry employment in the enterprise sector shifts toward industries subject to greater drops in U.S. tariffs on Vietnamese exports.[53] This evidence of shifting employment in industries with larger tariff cuts among employers in the enterprise sector (but not overall) is consistent with the framework in Bernard, Redding, and Schott (2007) as employers in the enterprise sector more directly benefit from lower variable exporting costs.[54]

[51] We use end of year firm-level employment in 2000 and 2003 and aggregate firm-level information to compute employment shares at the industry level. When we use end of year employment in 2001 instead of 2000, such that the timing more closely matches that of the VHLSSs, we obtain very similar results. The advantage of using 2000, however, is that we have a cleaner pre-BTA measure of employment.

[52] The results in Table 7 are robust to controlling for pre-existing trends in industry employment. These results are available upon request.

[53] The enterprise sector could grow because either workers are leaving household businesses for employers in the enterprise sector or because existing household businesses are formalizing and registering as private enterprises. Additional evidence suggests that the majority of workers move to the enterprise sector by finding a new job/employer in the enterprise sector. First, summary statistics suggests little mobility of household businesses to the enterprise sector. During this period, the number of registered private enterprises increased significantly, from about 35,000 in 2000 to about 84,000 in 2004 (Malesky and Taussig (2009)). Although this is an impressive growth in the number of private enterprises, it is only a small fraction of the estimated 8 million operating household businesses during this period (based on our own calculations from the household business modules of the 2002 and 2004 VHLSSs). Second, panel evidence on household businesses in Vietnam suggest that formalization is very rare. Between the 2006 and 2008 VHLSSs, only 1.7% of household businesses that continue to operate become registered as an enterprise (McCaig and Pavcnik (2017)). Consistent with this view, existing studies suggest little job creation in micro enterprises that formalize (see de Mel, McKenzie, and Woodruff (2013)).

[54] For example, products produced by firms in the household business sector might be poor substitutes for the products of firms in the enterprise sector.

The finding of employment reallocation across industries within the formal sector in panel B of Table 7 might seem at odds with the findings on trade policy and industry employment of several previous studies in the literature that use data that only covers employment in formal firms.[55] However, the difference in our results and those in the abovementioned studies can be reconciled by the difference in the type of trade liberalization studied. The earlier studies focus on unilateral trade liberalizations that reduced tariffs on imported goods entering the home market. We, on the other hand, focus on an episode of trade liberalization that primarily lowered tariffs in an export market. Recent theory on trade and firm heterogeneity highlights that the nature of trade reform matters for the predictions about the consequences of trade on the composition of employers/firms and their outcomes within an industry (see Melitz and Redding (2014) for a survey). Our setting and results are consistent with the predictions of these models, which predict reallocation of labor toward more productive firms in response to export market liberalization. Along those lines, our evidence is consistent with evidence on responses of other formal firm outcomes in studies that have examined the effects of export-market liberalization for firm technology, product quality, or skill upgrading among formal firms (Bustos (2011a, b), Verhoogen (2008), Brambilla, Lederman, and Porto (2012)).

More generally, our analysis highlights that the expansion of employment in the formal enterprise sector occurs, in part, through the reallocation of workers previously employed in household businesses. This demonstrates the importance of comprehensive micro data for exploring the various mechanisms of employment reallocation.

6. Worker Allocation and Aggregate Labor Productivity in Manufacturing

The reallocation of workers from household businesses to employers in the enterprise sector has potential implications for aggregate output.[56] In this section, we follow the macroeconomic

[55] Studies that use data covering employment in formal manufacturing firms above a certain employment cut-off include Revenga (1997), Currie and Harrison (1999), Wacziarg and Wallack (2004), and Topalova (2010). Topalova (2010) uses census and household survey data that are nationally representative for the district level analysis, while the analysis of reallocation across industries (Table 7, Panel A) is based only on the formal manufacturing sector.

[56] There is a large literature documenting the potential importance of firm- and sectoral-level distortions for aggregate output and productivity differences across countries. See, for example, Gollin (2008), La Porta and Shleifer (2008, 2014), Banerjee and Duflo (2005), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), McMillan and Rodrik (2011), and Gollin, Lagakos, and Waugh (2014).

development accounting literature to assess the potential impact of the BTA through this reallocation channel on aggregate labor productivity in manufacturing.

Consider a sector composed of two types of firms, household businesses and enterprises, which differ in their underlying labor productivity. A standard accounting formula evaluates the potential contribution to aggregate productivity stemming from the reallocation of labor across the two firm types as $\Delta P = s_e^{BTA} \Delta p_e^{BTA}$, where s_e^{BTA} is the share of manufacturing workers reallocated from the household business to the enterprise sector due to the BTA and Δp_e^{BTA} is the change in labor productivity for these workers as they reallocate. s_e^{BTA} can be computed using the estimated coefficient on tariffs from Table 3. A key issue is measurement of the productivity gap between the enterprise and household business sectors, which we turn to next.

6.1 Labor productivity gap

We first compute the labor productivity gap between firms in the enterprise and household business sectors using the common practice in the development accounting literature. Consider an industry composed of two types of firms, household businesses and enterprises, which differ in their underlying total factor productivity. Following Caselli (2005) and Gollin et al. (2014), we assume a Cobb-Douglas production function of the form $Y_s = A_s K_s^{\alpha_s} L_s^{1-\alpha_s}$, where A is total factor productivity, K is capital, L is labor, $(1-\alpha_s)$ is the output elasticity with respect to labor, and $s \in \{e,h\}$ denotes the enterprise and household business sectors respectively. If labor is homogeneous and markets are perfectly competitive, wages are equal to the marginal revenue product of labor (MRPL): $w_s = MRPL_s = (1-\alpha_s) ARPL_s$, where $ARPL_s$ is average revenue product of labor in sector s. This leads to the well-known result that with a Cobb-Douglas production function, the gap in the marginal revenue product of labor across the two sectors is proportional to the gap in the observed average revenue product of labor across the two sectors: $\frac{w_e}{w_h} = \frac{MRPL_e}{MRPL_h} = \frac{ARPL_e}{ARPL_h}$.[57]

This framework suggests two ways for calculating labor productivity gaps across the household business and enterprise sectors: wages and revenue per worker. Both of these measures have recently been used to compute productivity gaps between agriculture and non-agriculture, with

[57] Output elasticities of labor may also differ across sectors, which we return to later on.

Vollrath (2014) and Herrendorf and Schoellman (2015) using wages and Gollin et al. (2014) using revenue per worker. As in most other studies, the above approach computes revenue-based productivity assuming it is proportional to physical productivity. The revenue-based productivity gap might capture demand shocks or markup differences (see De Loecker, Goldberg, Khandelwal, and Pavcnik (2016)).[58]

Our data enables us to compute the productivity gap using both measures. For the enterprise sector, we compute the average revenue product of labor, ARPL, in the sector based on revenue per worker from firm-level data that covers all registered firms (the Vietnamese Enterprise Survey). For the household business sector, we compute ARPL based on aggregate revenue and the total number of workers from the household business and labor modules of the VHLSS. We compute the wage ratio using total annual earnings (this includes wage/salary payments as well as other payments such as public holiday payments and social allowance payments) among wage workers in the two sectors based on the labor module of the VHLSS. The details of these calculations are provided in Appendix B, Section B.1.1.

Table 8 reports the results. Row 1 of Table 8 reports the productivity gap based on ARPL in column 1 and wages in column 2. Both ratios exceed one, suggesting the possibility of aggregate productivity improvements from the reallocation of workers toward the higher productivity enterprise sector. However, the two ratios differ significantly. The ARPL ratio is 9.0, while the wage ratio is 1.82, almost 5 times smaller. Our estimate of the ARLP gap is consistent with large labor productivity gaps between informal and formal firms in other developing countries. For example, Nataraj (2011) reports that output per worker is 12.4 times higher in formal firms than informal firms in India.[59]

There are two issues with the above approach. First, the large labor productivity gap between the enterprise and household business sector in row 1 of Table 8 could in part reflect worker heterogeneity between the two sectors as, for example, workers in the enterprise sector are better educated on average. Second, the above framework cannot account for the difference in the magnitude of the labor productivity gap computed based on ARPL and wages. To the extent that firms in the household business sector face different distortions than firms in the enterprise sector,

[58] Analysis in De Loecker et al. (2016) requires firm-level prices. Like most studies, we do not have this information.

[59] La Porta and Shleifer (2014) report that value added per worker is 6.7 times higher in the formal than the informal sector in the median country in their sample. See Appendix B.1.2 for further discussion of the literature.

and these distortions create gaps between the payment received by a worker and the marginal revenue product of labor (Hsieh and Klenow (2009)), these differences could explain the disparities in the ARPL and wage gaps. We turn to these issues next.[60]

6.2 Interpreting the Labor Productivity Gap

6.2.1 Worker Heterogeneity

In the absence of data on the composition of workers, computation of a productivity gap using the ARPL from national accounts, industry-level, or firm-level data requires the assumption that labor is homogeneous across sectors. We relax this assumption and use additional worker-level information from the VHLSS to adjust the labor productivity gap for worker heterogeneity across the enterprise and household business sectors. As noted in Appendix Table A.3, workers in the enterprise sector work more hours annually and have higher levels of education than workers in the household business sector, implying that the productivity gap is overstated. We adjust the productivity ratios for differences in hours worked and human capital as in Gollin et al. (2014).

This adjustment significantly reduces the productivity gap, emphasizing the importance of accounting for worker heterogeneity across sectors. The results are presented in row 2 of Table 8. Consider the gap computed from wages in column 2. The gap drops by 0.58 from 1.82 to 1.24. In order for the wage gap to be fully eliminated, it would have to drop by 0.82 to 1. Worker heterogeneity therefore accounts for 70% of the wage gap (i.e., 0.58/(1.82-1)). Once we adjust for worker heterogeneity, the wage ratio is substantially closer to 1. Consider now the gap based on ARPL. As column 1 suggests, the ARPL ratio decreases from 9.0 to 6.0 after the adjustment. For the gap to be entirely eliminated it would have to decrease by 8 to 1. Thus, worker heterogeneity accounts for 37% of the labor productivity gap (3/(9-1)). This illustrates that accounting for worker heterogeneity matters.

[60] The above wage analysis excludes the self-employed, as they do not report a wage. Self-employment is uncommon in the enterprise sector: only 1.6% of manufacturing enterprise sector workers are self-employed (based on the 2004 VHLSS, as the 2002 VHLSS does not separately identify self-employment in the enterprise sector from self-employment in general). Hence, their omission from the wage ratio calculation is unlikely to significantly influence the results. However, the majority (61%) of manufacturing household business sector workers are self-employed in 2002. On average, the self-employed within the manufacturing household business sector have slightly more years of formal education (7.9 relative to 7.4) than wage workers within the sector and work a similar number of hours as wage workers (1851 vs 1948) in the household business sector. Hence, their unaccounted earnings would likely increase mean earnings in the household business sector and therefore decrease the wage ratio across the sectors. See further discussion and analysis in Appendix B.1.3.

ARPL and wage gaps continue to be present when we adjust for worker heterogeneity and additionally consider potential differences in industry composition and location of enterprises and household businesses in Table 8. For example, in columns 3 and 4 we report the two ratios for just one large manufacturing industry, textiles and apparel, so that differential industry composition across the two sectors does not influence the ratios. The adjusted ARPL and wage ratios in row two are 4.7 and 1.28, respectively. The gaps remain when computed for one major manufacturing area, the neighboring provinces of Ho Chi Minh City and Dong Nai (ARPL gap of 5.5 in column 5 and wage gap of 1.15 in column 6).[61]

The above approach controls for two dimensions of worker heterogeneity. Additionally, we estimate the wage gap for working in the enterprise sector by using Mincerian regressions, while controlling for worker heterogeneity in other dimensions, including location, gender, age, ethnic minority status, and industry affiliation. This analysis is discussed in detail in Section B.1.3 of Appendix B. After simultaneously controlling for these additional observable dimensions of worker heterogeneity (Appendix Table B.1), the wage gap is a similar order of magnitude as the adjusted wage gaps reported in Table 8. Relative to these estimates, using individual panel data and controlling for unobserved worker heterogeneity by including worker fixed effects reduces the hourly wage gap to 9% (Appendix Table B.2). Finally, the estimate of the earnings gap is a similar order of magnitude when we estimate Mincerian-style regressions that also include the self-employed and use hourly income as a dependent variable (Appendix Table B.3).

The bottom line that emerges from this analysis is that worker heterogeneity accounts for almost 40% of the original gap in ARPL across the two sectors and for 70% of the original gap in wages. Otherwise, a substantial part of the labor productivity gap simply reflects worker heterogeneity across sectors rather than labor productivity gaps for sectors with observationally equivalent workers.

6.2.2 Possible role of distortions

The adjusted ARPL gaps in Table 8 still substantially exceed the corresponding wage gaps. Our findings of a large ARPL ratio and small wage ratio is consistent with results on labor productivity gaps between agriculture and non-agriculture, namely large gaps in Gollin et al. (2014), which focuses

[61] These two provinces account for 21.2% of total manufacturing employment in the 2002 VHLSS and 38.3% of enterprise manufacturing employment in the 2001 enterprise data.

on average revenue product of labor; as compared to much smaller wage gaps in Vollrath (2014) and Herrendorf and Schoellman (2015). One potential explanation for the differences in ARPL and wage gaps are distortions imposed on firms as in Hsieh and Klenow (2009), which create gaps between the payment received by a worker and the marginal revenue product of labor. If firms in the household business sector face different distortions than firms in the enterprise sector, these differences could explain the differences in the ARPL versus wage gap.

Consider the framework in Hsieh and Klenow (2009), where firms in each sector face distortions in their profit function: $\pi_s = (1-\tau_{Ys})P_sY_s - (1+\tau_{Ls})w_sL_s - (1+\tau_{Ks})r_sK_s$, where τ_{Ys}, τ_{Ls}, and τ_{Ks} are the distortions in sector *s* for revenue, labor, and capital respectively. The optimal amount of labor in each sector s is found by setting $w_s = \frac{(1-\tau_{Ys})}{(1+\tau_{Ls})} MRPL_s = \frac{(1-\tau_{Ys})}{(1+\tau_{Ls})}(1-\alpha_s) ARPL_s$.

The average revenue product of labor across the two sectors can then be expressed as:

$$\frac{ARPL_e}{ARPL_h} = \frac{w_e}{w_h}\frac{(1-\alpha_h)}{(1-\alpha_e)}\frac{(1+\tau_{Le})}{(1+\tau_{Lh})}\frac{(1-\tau_{Yh})}{(1-\tau_{Ye})}. \quad (4)$$

The above expression illustrates that even if the ratio of wages is small, large differences in ARPL may persist due to a combination of differences in revenue distortions, labor market distortions, and output elasticities of labor. In this setting, which is more general than the one in Section 6.1, wage differences between sectors do not necessarily reflect overall marginal labor productivity differences across firms in the two sectors. Nonetheless, the wage gap is still useful as it captures the information on the potential income gap facing workers across the two sectors. Consequently, to compute the potential gain in aggregate labor productivity, we use the ARPL gap, adjusted for worker heterogeneity, as a measure of the labor productivity gap between the enterprise and household business sectors.[62]

6.2.3 Other considerations

We further examine the sensitivity of the ARPL estimate to measurement error that differs across the two sectors, and the possibility that the ARPL gap in part reflects differences in output elasticity of labor across sectors. These issues are presented in detail in Sections B.1.4-B.1.6 of Appendix B and briefly summarized below.

[62] See Trefler (2004), Bloom and Van Reenan (2007) and papers cited at the beginning of Section 6.

First, another potential concern is measurement error that differs across the two sectors. One concern is that the ARPL gap mainly reflects measurement error related to combining two different data sources. We use two different data sources to compute the ARPL gap because we are not aware of any surveys in Vietnam or elsewhere that are nationally representative and include formal and informal firms.[63] In fact, even data on informal firms alone is scarce. While measurement error is a concern, it is unlikely that most of the gap reflects measurement error due to two sources of data. First, note that similarly large productivity differences exist between informal and formal firms in La Porta and Shleifer (2008, 2014), which rely on surveys that capture both informal and formal firms. Second, we can compare annual earnings per formal sector manufacturing worker between the 2002 VHLSS and 2001 enterprise data. Mean annual earnings per worker were 11.6 million VND in the 2002 VHLSS as compared to 11.8 million VND in the 2001 enterprise data. The similarity of these estimates suggests that survey differences do not necessarily lead to divergent responses. Third, as an additional check, we use the business module of a more recent household survey, the 2006 VHLSS, which distinguished between household businesses and private enterprises (but did not cover other firms in the enterprise sector, such as state-owned and foreign enterprises), and asked about business revenue in an identical manner to the 2002 VHLSS. We estimate an ARPL ratio of 3.6 (2.8 adjusted for worker heterogeneity) between private enterprises and household businesses (Appendix Table B.4, column 1). This comparison excludes state owned enterprises and foreign invested firms, and the private enterprises are less productive, on average, than state and foreign enterprises. Thus, we would expect, and indeed do find, a lower ARPL gap when household businesses are compared only to private enterprises. Nevertheless, an APRL gap remains in this comparison based on a single data source.

More generally, we evaluate the potential impact of measurement error in revenue or employment in the household business sector for ARPL gaps. Very few microenterprise operators keep formal accounts and thus measurement error is potentially more likely to affect our estimates of the ARPL in the household business sector than in the enterprise sector. In Appendix Section B.1.4 (and Appendix Table B.5), we adjust the ARPL gap for potential measurement error in reported revenue and in reported labor supply in household businesses, relying on estimates of possible measurement error based on detailed data on microenterprises from de Mel, McKenzie, and

[63] Nataraj (2011), Hsieh and Olken (2014), and Ulyssea (2017) also use informal and formal firm data from two datasets.

Woodruff (2009) and Fafchamps et al. (2014). For example, de Mel, McKenzie, and Woodruff (2009) suggest that reported revenue in microenterprises may underestimate true revenue by as much as 30%. When we adjust the reported revenue in the household business sector by this factor, the labor productivity gap falls from 6 to 4.2. The other concern is that people working in the household business sector may overstate effective hours worked. Data collected by Fafchamps et al. (2014) reports information from Ghana on total hours worked and total hours worked with full effort, and suggest that microenterprise owners report working at full effort 89 percent of the time. Thus, reported hours worked may slightly overestimate true labor input.[64] When we adjust the productivity gap for potential measurement error in revenue and hours worked, the ARPL gap is 3.7. We focus on this productivity gap and report it in column 1 and row 3 of Table 8 because it is the most conservative measure from Appendix Table B.5. The above discussion suggests that the labor productivity gap between workers in the enterprise and household business sectors could partially reflect measurement error.

Second, equation (4) illustrates that the ARPL gap could in part reflect a lower output elasticity of labor in the enterprise sector within a given industry and not just the gaps in MRPL. Like much of the productivity gap literature, we so far assumed equal output elasticity of labor across heterogeneous sectors or firms. To be conservative, we also consider an alternative case using estimates from existing literature (see section B.1.6 of Appendix B and Appendix Table B.5 for details). For example, Restrepo-Echavarria (2014) assumes output-labor elasticities of 1 and 0.68 in the household business sector and enterprise sector, respectively, for a ratio of about 1.5.[65] This alternative case, where the informal sector uses no capital and only labor for production, yields a MRPL gap of 4. Note that the output-labor elasticity difference would have to be 6 to fully account for the productivity gap that adjusts for worker heterogeneity reported in column 1 of Table 8. When we adjust the labor productivity gap for worker heterogeneity, measurement error in revenue and hours worked, and differences in output- labor elasticities, it drops to 2.5. This gap is reported in row 4 of column 1 in Table 8.

Overall, the above analysis highlights the importance of considering worker heterogeneity, potential measurement error issues, and assumptions about output elasticity of labor in this

[64] Additionally, measurement issues specific to the 2002 VHLSS are discussed in Section B.1.5.
[65] Restrepo-Echavarria (2014) reports value added-labor elasticity. Adjusting for factor share of materials based on Nataraj (2011) yields similar results because the factor share of materials is very similar across the two sectors.

literature. The underlying assumption for the different output-labor elasticities across the two sectors assumed that the informal sector uses no capital and only labor for production, a very conservative assumption. As a result, we use the estimate of the labor productivity gap that adjusts for worker heterogeneity and measurement error in revenue and hours worked (i.e., 3.7) as our preferred estimate of the ARPL gap. We use the estimate that adjusts only for worker heterogeneity (i.e., 6) as the upper bound, and the estimate that adjusts for worker heterogeneity, measurement error in revenue and hours worked, and differences in output-labor elasticities (i.e., 2.5) as the lower bound.

6.3 Aggregate labor productivity in manufacturing

We use our estimates of the ARPL gap across sectors in Table 8 to calculate the potential gain in aggregate productivity within manufacturing in response to BTA-induced reallocation of workers from the household business to the enterprise sector. We evaluate the potential contribution of reallocation to aggregate productivity stemming from the reallocation of labor across the two sectors by using a standard development accounting formula introduced at the beginning of the section and expressing it as the percentage change in aggregate labor productivity, relative to the baseline aggregate labor productivity:

$$\frac{s^{BTA}\left(ARPLratio-1\right)ARPL_h}{\left(1-s_h\right)ARPL_e+s_hARPL_h}, \tag{5}$$

where s^{BTA} is the share of manufacturing workers reallocated from the household business to the enterprise sector due to the BTA, $ARPLratio$ is the ARPL gap , $ARPL_e$ and $ARPL_h$ are the initial average revenue per hour worked in the enterprise and household business sectors, respectively, and s_h is the initial share of hours worked in the household business sector.

The coefficient on the industry tariff in column 3 of Table 3 implies that the BTA reallocated 5.0 percent of manufacturing hours from household businesses to enterprises by 2003/04 (see Section B.2 of Appendix B for details). This is our measure of s^{BTA} . Based on our preferred estimate of the ARPL gap across sectors, which adjusts for worker heterogeneity and measurement error in revenue and hours worked, reported in row 3 of column 1 in Table 8, we find that the BTA-induced movement of workers increased ARPL per hour worked by 2.8 percent annually within manufacturing. This estimate focuses on productivity per hour worked. Since workers in the

enterprise sector work approximately 25 percent more hours annually, predicted growth in productivity per worker is 3.4 percent annually (see Section B.3 of Appendix B and Appendix Table B.6). This estimate of the annual aggregate labor productivity gains based on the ARPL gap is relatively large. As discussed in Section 5.5, the reallocation of workers is predominantly due to the creation of new jobs in enterprise sector firms as opposed to existing household businesses transitioning to the enterprise sector. Moreover, Table 8 illustrates that, depending on the assumptions about the measurement error issues and differences in output-labor elasticities between the sectors, the estimated aggregate gains in labor productivity could range from 3.5 (when we assume no measurement error differences between the two sectors) to 1.5 percent per year (when we adjust for measurement error differences and allow for a large output elasticity of labor difference across the two sectors). Appendix section B.3 discusses the alternative estimates in greater detail.

We use a similar formula as above to compute the associated gains in hourly wages for workers due to reallocation, which is 0.5 percent per year (Table 8, column 2). The gains in annual wages for workers are slightly larger, 0.9 percent per year (see Appendix Table B.6, column 2), when one also takes into account the difference in hours worked between the two sectors on an annual basis. The estimates based on the 9% wage premium for working in the enterprise sector among panel workers suggest gains in hourly wages at a rate of 0.19 percent per year.

In sum, our preferred estimate of the labor productivity gap suggests that the reallocation of labor from the informal to the formal sector in response to the BTA increased aggregate labor productivity within manufacturing by 2.8 percent per year in the two years following the BTA.

7. Conclusion

Vietnam's trade agreement with the U.S. provides an excellent setting to examine how declines in export costs affect the reallocation of employment across employers in a low-income country, where a majority of workers are employed in informal microenterprises. We find that the reallocation of labor from microenterprises to formal employers provides an important margin of adjustment to new exporting opportunities. Industries with bigger declines in export costs experience a greater reduction in household business employment, with workers in more internationally integrated provinces and in younger cohorts responding more strongly. Our results complement the existing literature on trade and labor reallocation in developing countries, which

has primarily focused on the effects of domestic import liberalization on the reallocation of workers across firms within the formal sector or across industries. Our estimates reflect short run responses and may underestimate the long run effects of the BTA on employment in the formal sector, as the economy has more time to adjust (Dix-Carneiro and Kovak (2017)).

The difference in the type of trade reform may help explain why our findings diverge from the literature that found no formal sector employment increase in the short-run after import tariff liberalization in developing countries. Factors such as differences in mobility of labor across regions could also pay a role. Overall, further exploration of the relationship between the type of trade reform, imperfections in the domestic product or factor markets and worker outcomes, and longer-run responses to the trade shock remains a fruitful area for future research. In addition, our findings might generalize to other low-income country settings as they are robust to using self-employment, a definition of informal microenterprise employment that does not depend on a country-specific legal definition of informality, as a dependent variable. They are more likely to generalize to other low-income countries with a comparative advantage in low-skill manufacturing where production takes place in both informal microenterprises and larger, formal firms.

The movement into a formal sector firm has potentially important consequences for workers. We show that working in a formal enterprise changes the way a worker is attached to the workforce. In Vietnam, workers in the enterprise sector earn higher wages, are more likely to receive non-wage/salary payments, work longer hours, and are less likely to hold multiple jobs. At the same time, our analysis highlights that it is crucial to take into account worker heterogeneity and sorting in assessing wage and earnings differences across the two sectors.

Our results also relate to the literature that emphasizes the implications of the inefficient allocation of resources across heterogeneous firms for aggregate productivity (Restuccia and Rogerson (2008), Hsieh and Klenow (2009)). We show that the removal of an output market distortion that is more binding for initially more productive firms, such as a tariff on exports, induces a movement of workers away from less-productive employers in informal microenterprises to employers in the more productive enterprise sector. Due to firm data constraints, one usually cannot observe the entire distribution of firms, both informal and formal, in low-income countries. The use of labor force data provides an alternative for observing the allocation of workers across this margin of the firm distribution in response to trade if the labor force data includes information

on employer type and informal sector firm data is not available (see Dix-Carneiro and Kovak (2017) for a recent example).

We also evaluate the labor productivity gap between the formal and informal sectors within manufacturing using detailed micro-survey data from nationally representative surveys of both informal microenterprises and formal firms. Our preferred estimate of the labor productivity gap of 3.7 suggests that the BTA increased aggregate labor productivity by 2.8 percent annually in the two years following the BTA due to moving labor from the informal to the formal sector. Our analysis highlights some of the challenges in estimating this productivity gap, including the importance of taking into account worker heterogeneity, potential measurement error issues that might be particularly large in the informal sector, and differences in output-labor elasticities across the two sectors. Adjustment for all these issues substantially reduces the estimates of the labor productivity gap and the implied change in labor productivity, lowering it to only 1.5 percent per year. More broadly, this finding has implications for the reallocation and misallocation literature, which uses the gap to evaluate the potential aggregate productivity gains from the reallocation of workers or the removal of distortions. To the extent that used measures of distortions do not account for worker heterogeneity, such exercises might overestimate the aggregate productivity gains from the removal of distortions.

Given the prevalence of informal microenterprises in low-income countries and increasing availability of better micro survey data, we expect that studying the determinants of the prevalence of informal microenterprises and the sources of the labor productivity gap between informal and formal firms will continue to be a topic for future research.

References

Atkin, David. (2009). "Working for the future: Female factory work and child health in Mexico," unpublished manuscript.

Attanasio, O., P. Goldberg, and N. Pavcnik. (2004). "Trade reforms and wage inequality in Colombia," *Journal of Development Economics*, 74, 331-366.

Banerjee, Abhijit V. 2013. "Microcredit under the microscope: What have we learned in the past two decades, and what do we need to know?" *Annual Review of Economics*, 5, p. 487-519.

Banerjee, A. and E. Duflo (2005). "Growth Theory through the Lens of Development Economics," in eds. Aghion, P. and S. Durlauf. Handbook of Economic Growth, North Holland.

Banerjee, Abhijit V. and Esther Duflo. (2007). "The Economic Lives of the Poor" *Journal of Economic Perspectives*, 21(1), pp. 141-168.

Banerjee, Abhijit V. and Esther Duflo. (2008). "What is middle class about the middle classes around the world?" *Journal of Economic Perspectives*, 22(2), pp. 3-28.

Banerjee, Abhijit V. and Esther Duflo. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*, PublicAffairs: New York City.

Banerjee, Abhijit V., Dean Karlan, and Jonathan Zinman. 2015. "Six randomized evaluations of microcredit: Introduction and further steps," *American Economic Journal: Applied Economics*, 7(1), p. 1-21.

Bento, P. and D. Restuccia. (2016). "Misallocation, establishment size, and productivity," University of Toronto Department of Economics Working Paper 557.

Bernard, A., S. Redding, and P. Schott (2007). "Comparative advantage and heterogeneous firms," *Review of Economics Studies*, Vol. 74, No. 1, pp. 31-66.

Bloom, Nick and John Van Reenen. (2007). "Measuring and explaining management practices across firms and countries," *The Quarterly Journal of Economics*, 122(4), pp. 1351-1408.

Bosch, M., E. Goñi-Pacchioni, and W. Maloney. (2012). "Trade liberalization, labor reforms and formal–informal employment dynamics," *Labour Economics*, 19(5), 653-67.

Brambilla, I., D. Lederman and G. Porto (2012). "Exports, Export Destinations, and Skills," *American Economic Review*, Vol. 102, No. 7, pp. 3406-3438.

Brambilla, I., G. Porto, and A. Tarozzi (2012). "Adjusting to trade policy: Evidence from U.S. antidumping duties on Vietnamese catfish," *Review of Economics and Statistics*, Vol. 94, No. 1, pp. 304-319.

Bruhn, Miriam and David McKenzie. 2014. “Entry regulation and the formalization of microenterprises in developing countries,” *The World Bank Research Observer*, 29(2), p. 186-201.

Bustos, P. (2011a). "Trade liberalization, exports, and technology upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms," *American Economic Review*, 101(1), 304-40.

Bustos, P. (2011b). “The Impact of Trade Liberalization on Skill Upgrading: Evidence from Argentina,” unpublished manuscript.

Caselli, F. (2005). “Accounting for Cross-Country Income Differences,” in eds. Aghion, P. and S. Durlauf. Handbook of Economic Growth, North Holland.

Coşar, A. K. (2013). “Adjusting to trade liberalization: Reallocation and Labor Market Policies,” mimeo.

Currie, J. and A. Harrison. (1997). “Sharing the costs: The impact of trade reform on capital and labor in Morocco,” *Journal of Labor Economics*, 15(3), S44-71.

Davis, Steven J. and John Haltiwanger. (1992). “Gross job creation, gross job destruction, and employment reallocation,” *The Quarterly Journal of Economics*, 107(3), pp. 819-863.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik. (2016). “Prices, markups, and trade reform,” *Econometrica*, 84(2), 445-510.

de Mel, Suresh, David J. McKenzie, and Chris Woodruff. 2009. “Measuring microenterprise profits: Must we ask how the sausage is made?” *Journal of Development Economics*, 88(1), 19-31.

de Mel, Suresh, David McKenzie, and Christopher Woodruff (2013). "The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka." *American Economic Journal: Applied Economics*, 5(2): 122-50.

Demidova, S. and A. Rodríguez-Clare. (2013). “The simple analytics of the Melitz model in a small open economy,” *Journal of International Economics*, 90(2), 266-72.

Dix-Carneiro, Rafael. (2014). “Trade liberalization and labor market dynamics”, *Econometrica*, 82(3), pp. 825-885.

Dix-Carneiro, Rafael and Brian Kovak. (2017). “Trade liberalization and regional dynamics,” *American Economic Review*, 107(10), pp. 2908-46.

Fafchamps, Marcel, David McKenzie, Simon Quinn, and Christopher Woodruff. (2014). “Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in Ghana,” *Journal of Development Economics*, 106, 211-226.

Feliciano, Zadia M. (2001). “Workers and trade liberalization: The impact of trade reforms in Mexico on wages and employment,” *Industrial and Labor Relations Review*, 55(1), pp. 95-115.

Fernández, Andrés and Felipe Meza. (2015). “Informal employment and business cycles in emerging economies: The case of Mexico,” *Review of Economic Dynamics*, 18, pp. 381-405.

Fukase, Emiko. (2013). “Export liberalization, job creation and the skill premium: Evidence from the U.S.-Vietnam Bilateral Trade Agreement,” *World Development*, 41, 317-337.

Gollin, Douglas. (2002). "Getting Income Shares Right," *Journal of Political Economy*, 110 (2), pp.458-474.

Gollin, D. (2008). “Nobody’s business by my own: Self-employment and small enterprise in economic development,” *Journal of Monetary Economics* 55, 219-233.

Gollin, D., D. Lagakos, and M. Waugh (2014). “The Agricultural Productivity Gap in Developing Countries,” *Quarterly Journal of Economics*, 129(2), 939-993.

Goldberg, P. and N. Pavcnik (2003). “The Response of the Informal Sector to Trade Liberalization”, *Journal of Development Economics,* 72, 463-496.

Goldberg, P. and N. Pavcnik (2007). "Distributional Effects of Globalization in Developing Countries", *Journal of Economic Literature* March 2007, 45(1), pp. 39-82

Hanson, G. (1996). “Economic Integration, Intraindustry Trade, and Frontier Regions,” *European Economic Review* 40(3-5), pp. 941-949.

Hanson, G. (2012). “The rise of Middle Kingdoms: Emerging economies in global trade,” *Journal of Economic Perspectives*, Vol. 26, No. 2, pp. 41-64.

Harrison, A., J. McLaren, and M. McMillan (2011). “Recent perspectives on trade and inequality,” *Annual Review of Economics*, Vol. 3, pp. 261-289.

Heath, Rachel and A. Mushfiq Mobarak. (2015). “Manufacturing growth and the lives of Bangladeshi women,” *Journal of Development Economics*, 115, pp. 1-15.

Herrendorf, Berthold and Todd Schoellman. (2015). “Wages, human capital, and structural transformation,” draft.

Hsieh, Chang-Tai and Peter J. Klenow (2009). “Misallocation and manufacturing TFP in China and India,” *Quarterly Journal of Economics*, Vol. 124, No. 4, pp. 1403-1448.

Hsieh, C.-T. and B. Olken. (2014). “The missing ‘Missing Middle’,” *Journal of Economic Perspectives*, 28, 89-108.

Javorcik, Beata S. (2015). “Does FDI bring good jobs to host countries?” *The World Bank Research Observer*, 30(1), pp. 74-94.

Jensen, Robert. (2012). “Do labor market opportunities affect young women’s work and family decisions? Experimental evidence from India,” *The Quarterly Journal of Economics*, 127(2), pp. 753-792.

La Porta, R. and A. Shleifer. (2008). “The unofficial economy and economic development,” *Brookings Papers on Economic Activity*, 39, 275-363.

La Porta, R. and A. Shleifer. (2014). “Informality and development,” *Journal of Economic Perspectives*, 28, 109-26.

Lucas, R. (1978). “On the size distribution of business firms,” *The Bell Journal of Economics*, 9, 508-23.

Kovak, B. (2013). “Regional effects of trade reform: What is the correct measure of liberalization?” *American Economic Review*, 103(5), 1960-76.

Marcouiller, D., V. Ruiz de Castilla, and C. Woodruff (1997): “Formal Measures of the Informal-Sector Wage Gap in Mexico, El Salvador, and Peru," *Economic Development and Cultural Change,* 45:2, pp. 367-392.

Malesky, E. and M. Taussig (2009). “Out of the gray: The impact of provincial institutions on business formalization in Vietnam,” *Journal of East Asian Studies*, Vol. 9, pp. 249-290.

McCaig, Brian (2011). “Exporting out of poverty: Provincial poverty in Vietnam and U.S. market access.” *Journal of International Economics*, 85(1), pp. 102-113.

McCaig, Brian and Nina Pavcnik. (2015). “Informal employment in a growing and globalizing low-income country,” *American Economic Review Papers and Proceedings*, 105(5), 545-550.

McCaig, Brian and Nina Pavcnik. 2017. “Out with the old and unproductive, in with the new and similarly unproductive: Microenterprise dynamics in a growing low-income country." GLM|LIC Working Paper No. 23.

McKenzie, David and Christopher Woodruff. 2014. “What are we learning from business training and entrepreneurship evaluations around the developing world?” *The World Bank Research Observer*, 29(1), p. 48-82.

McMillan, M. and D. Rodrik. (2011). “Globalization, structural change and productivity growth,” NBER Working Paper No. 17143.

Melitz, M. (2003): “The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity,” *Econometrica* 71, 1696-1725.

Melitz, M. and S. J. Redding. (2014). “Heterogeneous firms and trade,” in Gopinath, G., E. Helpman, and K. Rogoff. (eds). Handbook of International Economics, Volume 4: Elsevier.

Menezes-Filho, N. and M. Muendler (2011). "Labor Reallocation in Response to Trade Reform," *NBER Working Paper* 17372.

Mrazova, M and P. Neary (forthcoming). “Selection Effects with Heterogeneous Firm,” *Journal of European Economic Association*.

Nataraj, S. (2011). “The impact of trade liberalization on productivity: Evidence from India’s formal and informal manufacturing sectors,” *Journal of International Economics* 85, 292-301.

Paz, L. (2014). “The impacts of trade liberalization on informal labor markets: A theoretical and empirical evaluation of the Brazilian case,” *Journal of International Economics*, 92, 330-348.

Pregelj, V. (2005). “Normal-Trade-Relations (Most-Favored-Nation) Policy of the United States,” Congressional Research Service Report RL31558.

Restrepo-Echavarria. (2014). “Macroeconomic volatility: The role of the informal economy,” *European Economic Review*, 70, pp. 454-469.

Restuccia, D. and R. Rogerson. (2008). “Policy distortions and aggregate productivity with heterogeneous establishments,” *Review of Economic Dynamics*, 11, 707-720.

Restuccia, D., D. Yang, and X. Zhu. (2008). “Agriculture and aggregate productivity: A quantitative cross-country analysis,” *Journal of Monetary Economics*, 55, 234-250.

Revenga, A. (1997). “Employment and wages effects of trade liberalization: The case of Mexican manufacturing,” *Journal of Labor Economics*, 15(3), S20-43.

STAR-Vietnam. (2003). “An Assessment of the Economic Impact of the United States – Vietnam Bilateral Trade Agreement.” The National Political Publishing House: Hanoi, Vietnam.

Topalova, P. (2010). “Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India,” *American Economic Journal: Applied Economics*. Vol. 2, No. 4, pp. 1-41.

Trefler, D. (2004). “The long and short of the Canada-U.S. Free Trade Agreement,” *American Economic Review*, 94(4), 870-95.

Tybout, J. R. (2000). “Manufacturing firms in developing countries: How well do they do, and why?” *Journal of Economic Literature*, 38(1), 11-44.

Tybout, J. R. (2014). “The missing middle: Correspondence,” *Journal of Economic Perspectives*, 28(4), 235-36.

Ulyssea, G. (2017). “Firms, informality and development,” unpublished manuscript.

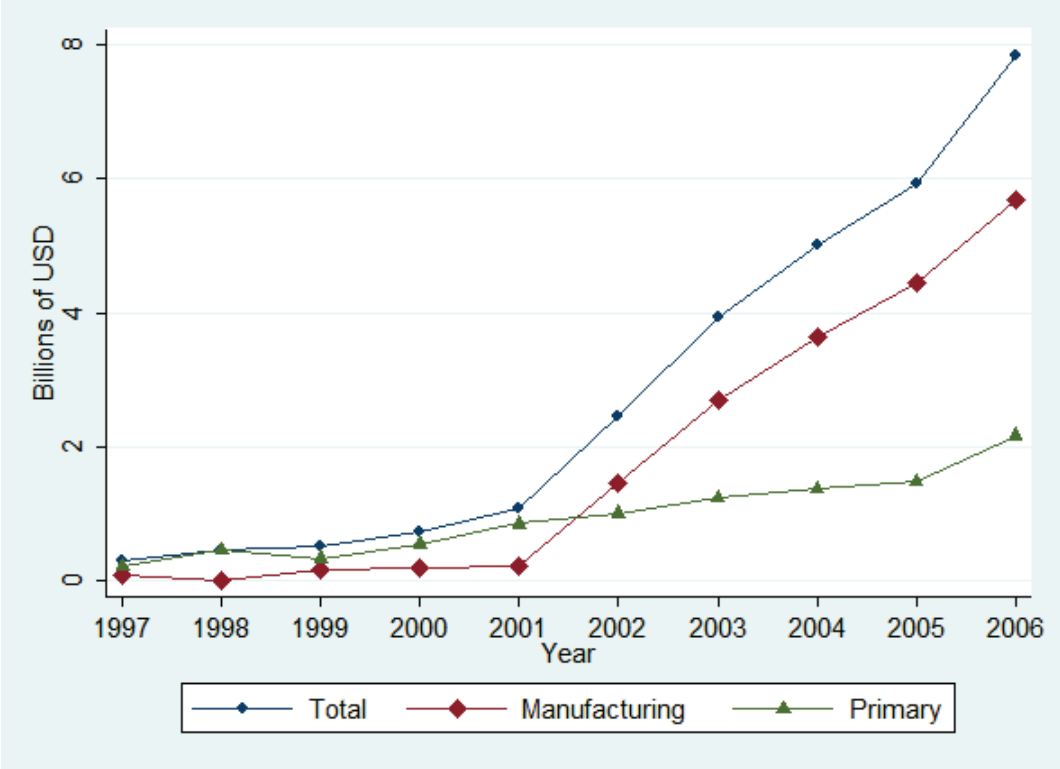
Verhoogen, E. (2008). "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, vol. 123, no.2, pp. 489-530.

Vollrath, Dietrich. (2014). "The efficiency of human capital allocations in developing countries," *Journal of Development Economics*, 108, pp. 106-118.

Wacziarg R. and J. Wallack (2004). "Trade liberalization and intersectoral labor movements," *Journal of International Economics*, 64(2), 411-439.

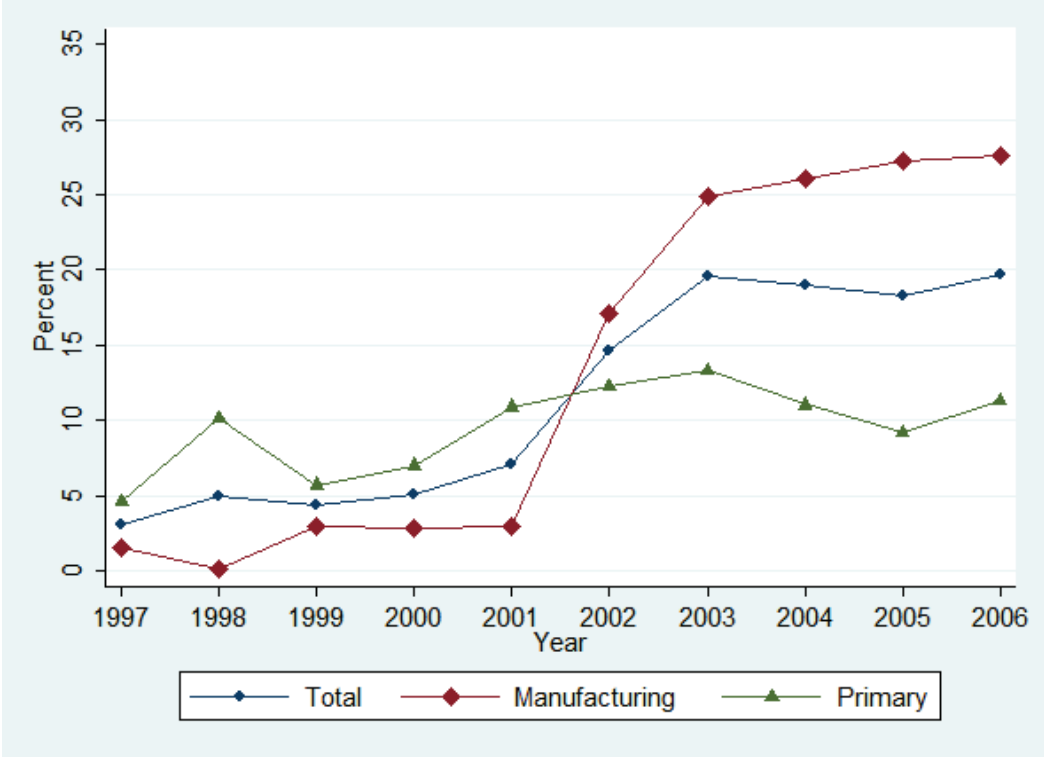
World Bank. 2011. *Vietnam Development Report 2012: Market Economy for a Middle-Income Vietnam*. Washington, DC: World Bank.

Figure 1: Value of Vietnamese exports to the U.S., 1997 to 2006



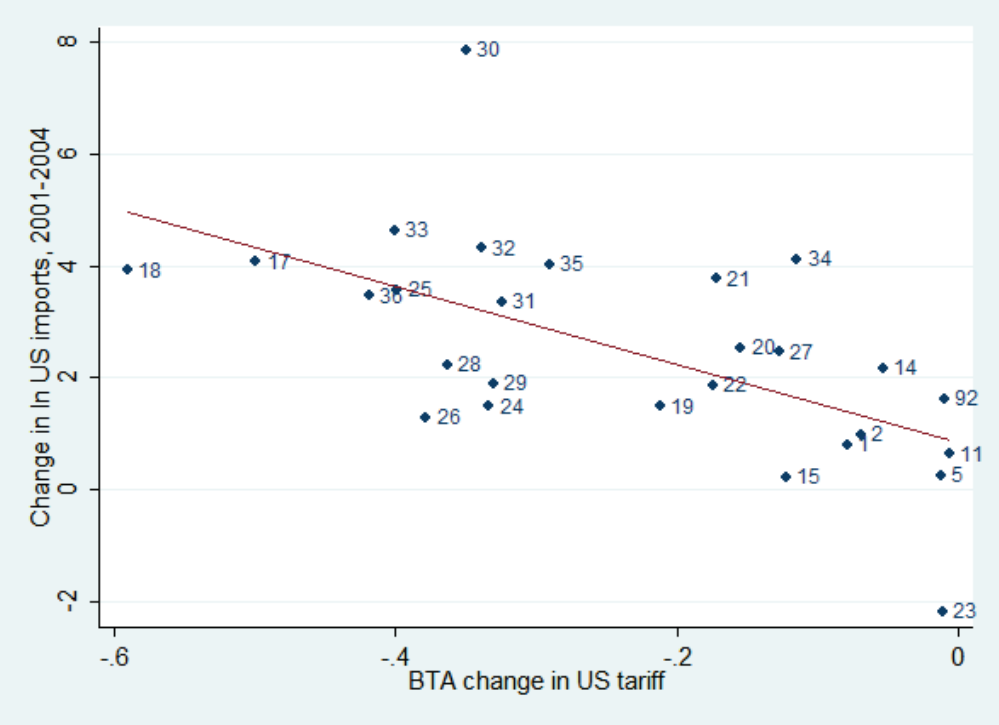
Notes: Authors' calculations from COMTRADE.

Figure 2: Share of the United States in Vietnam's Exports



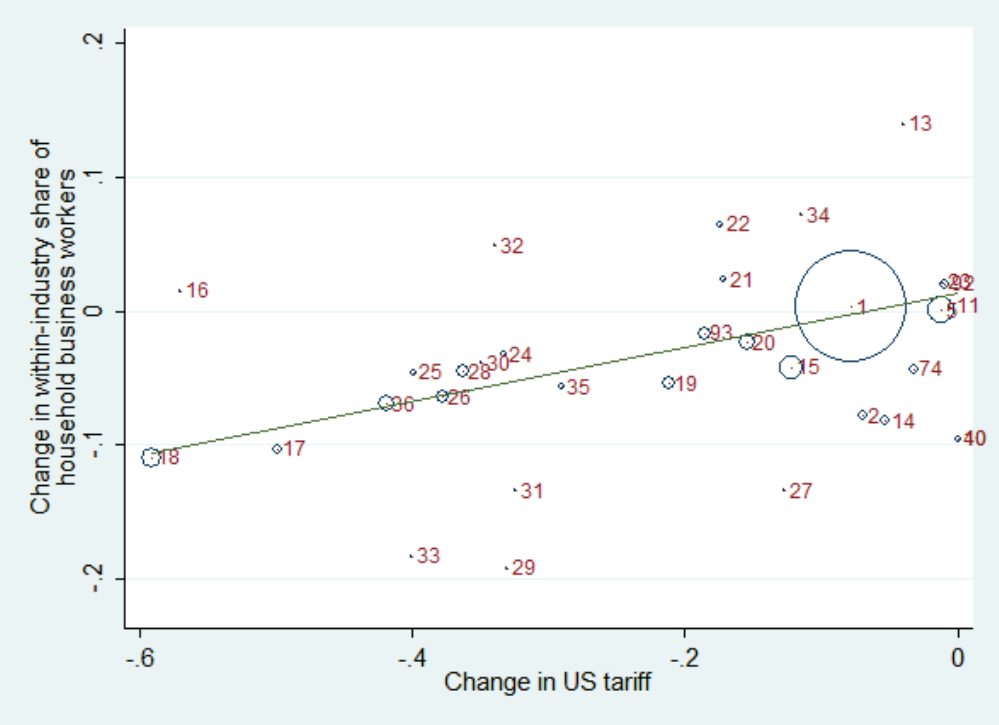
Notes: Authors' calculations from COMTRADE.

Figure 3: Growth of Vietnamese exports to the US versus US tariff cuts by industry



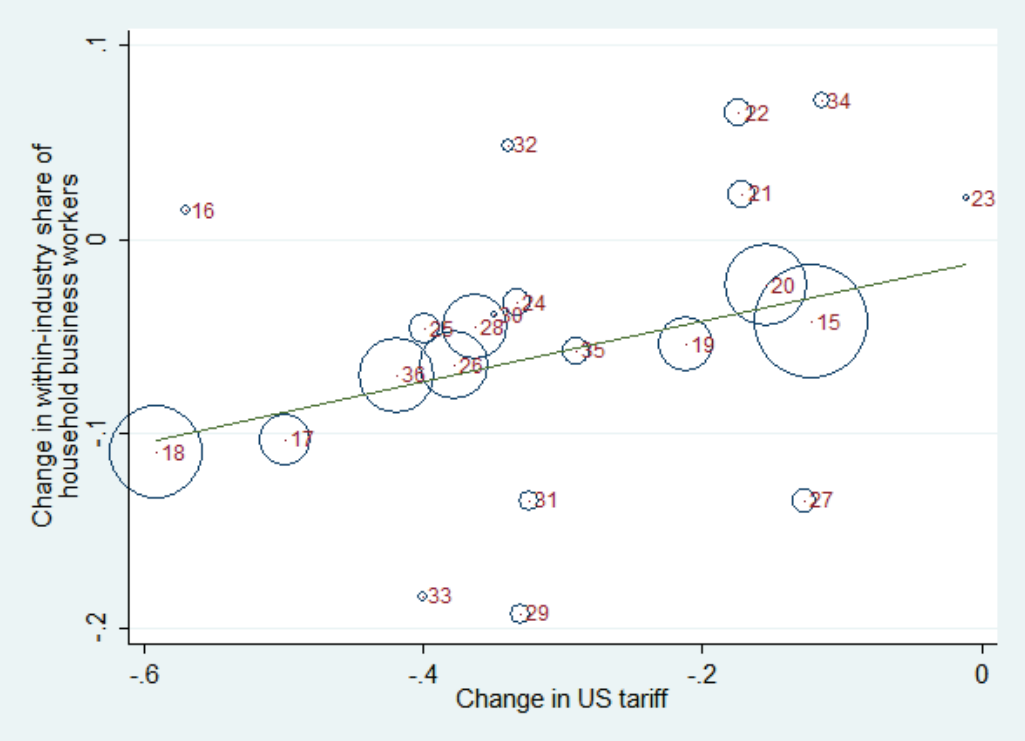
Notes: The industry codes correspond to ISIC revision 3.

Figure 4: Change in share of household business workers and U.S. tariff reductions



Notes: The bubble sizes represent the weight given to the industry in the plotted regression line. See text for explanation. The industry codes correspond to all traded industries in ISIC revision 3.

Figure 5: Change in share of household business workers and U.S. tariff reductions, manufacturing industries



Notes: The bubble sizes represent the weight given to the industry in the plotted regression line. See text for explanation. The industry codes correspond to ISIC revision 3.

Table 1: Summary of U.S. tariffs applied to imports from Vietnam

Industry	Number of industries	Mean pre-BTA tariff (Column 2)	Mean post-BTA tariff (MFN)	Mean change in tariff	Standard deviation of tariff change
Traded industries	34	0.234	0.025	-0.209	0.179
All industries	60	0.133	0.014	-0.119	0.170
Manufacturing	22	0.338	0.036	-0.302	0.153

Notes: The tariffs reported are simple averages across the indicated set of industries. Non-traded industries, which are included in "All industries" have been assigned a tariff of 0 both before and after the BTA.

Table 2: Share of employment in household businesses

	All	Excluding agriculture and fisheries	Manufacturing
Panel A: Share of employment in household businesses			
2002	0.847	0.672	0.656
2004	0.814	0.626	0.600
Panel B: Decomposing changes in household business employment			
Within industries	-0.017	-0.040	-0.059
Between industries	-0.016	-0.006	0.003
Total	-0.033	-0.046	-0.056

Notes: Authors' own estimates based on the 2002 and 2004 VHLSSs. Based on workers aged 20 to 64 inclusive. Survey sampling weights included.

Table 3: Employment in Household Businesses and Tariffs

Dependent variable: Indicator for working in a household business

	(1) Traded	(2) All	(3) Manufacturing
Panel A: Baseline specification			
Tariff	0.209***	0.127***	0.156***
	(0.0144)	(0.0323)	(0.0197)
R-squared	0.419	0.594	0.299
Panel B: Province-year fixed effects			
Tariff	0.212***	0.123***	0.170***
	(0.0195)	(0.0300)	(0.0277)
R-squared	0.420	0.594	0.304
Panel C: Province-year fixed effects and time varying individual covariates			
Tariff	0.203***	0.115***	0.152***
	(0.0232)	(0.0266)	(0.0266)
R-squared	0.420	0.595	0.305
Number of industries	34	60	22
Observations	176,544	248,791	27,072

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The sample is restricted to workers between the ages of 20 and 64 inclusive at the time of the survey. Column (1) includes all traded industries, column (2) includes all industries, and column (3) includes all traded manufacturing industries. All regressions include worker characteristics (age, age squared, education level indicators, female indicator, ethnic minority indicator, and rural indicator), as well as industry fixed effects. In Panel A we include province and year fixed effects. In Panel B, the province and year fixed effects are replaced by province-year fixed effects. In Panel C, all control variables, except industry fixed effects, are interacted with a 2004 indicator.

Table 4: Results and falsification test based on self-employment
Dependent variable: Indicator for self-employment

	(1) Traded	(2) All	(3) Manufacturing
Panel A: Reform period			
Tariff	0.212***	0.131***	0.200***
	(0.0405)	(0.0390)	(0.0501)
Observations	176,544	248,791	27,072
R-squared	0.370	0.532	0.240
Panel B: Pre-reform period (1993-1998)			
Tariff	0.0379	0.0157	-0.00870
	(0.0334)	(0.0357)	(0.0918)
Observations	17,559	22,753	2,367
R-squared	0.344	0.501	0.236

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The dependent variable is an indicator for whether an individual is self-employed. In Panel B, the pre-BTA tariffs (Column 2 rates in 2001) are assigned to industries in 1993 and the post-BTA tariff (MFN rates in 2001) are assigned to industries in 1998. All regressions include worker characteristics (age, age squared, education level indicators, female indicator, ethnic minority indicator, and rural indicator), as well as industry, province and year fixed effects.

Table 5: Employment in Household Business and Tariffs by Age, Gender, Education, and Location
Dependent variable: Indicator for working in a household business

	(1) Traded	(2) All	(3) Manufacturing
Panel A: Location			
Less than the median distance from a major	0.227***	0.142***	0.185***
seaport	(0.0215)	(0.0333)	(0.0269)
Observations	83,079	126,340	18,926
At least the median distance from a major	0.147***	0.0705*	0.0493
seaport	(0.0345)	(0.0365)	(0.0482)
Observations	93,465	122,451	8,146
Panel B: Age			
Age 20 to 29	0.327***	0.223***	0.175***
	(0.0512)	(0.0424)	(0.0610)
Observations	50,069	68,237	10,726
Age 30 to 39	0.142***	0.0821***	0.0808***
	(0.0198)	(0.0281)	(0.0240)
Observations	52,620	75,901	8,273
Age 40 to 49	0.119***	0.0379	0.138***
	(0.0276)	(0.0389)	(0.0349)
Observations	43,859	64,995	5,681
Age 50 to 59	0.107	0.0357	0.131
	(0.0648)	(0.0637)	(0.0914)
Observations	22,530	30,611	1,982
Age 60 to 64	-0.0186	-0.0658	-0.141
	(0.107)	(0.0849)	(0.173)
Observations	7,466	9,047	410
Panel C: Gender			
Males	0.229***	0.107**	0.154**
	(0.0384)	(0.0509)	(0.0610)
Observations	84,522	123,164	13,409
Females	0.196***	0.148***	0.159***
	(0.0178)	(0.0204)	(0.0317)
Observations	92,022	125,627	13,663
Panel D: Education			
Did not complete lower secondary	0.217***	0.146***	0.151***
	(0.0180)	(0.0406)	(0.0286)
Observations	106,125	132,298	11,193
Completed lower secondary	0.176***	0.0991***	0.126***
	(0.0168)	(0.0369)	(0.0292)
Observations	50,531	71,256	9,208
Completed upper secondary	0.201***	0.125***	0.193***
	(0.0358)	(0.0332)	(0.0612)
Observations	19,888	45,237	6,671

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The table shows the estimated coefficient on industry tariffs from regressing an indicator for working in a household business for the indicated sample. All regressions include the usual controls for worker characteristics, and province, industry, and year fixed effects as in Table 3 Panel A.

Table 6: Household business employment and tariffs, panel-level analysis
Dependent variable: Indicator for working in a household business

	(1) Traded	(2) All	(3) Manufacturing
Panel A: Tariff based on contemporary industry			
Tariff	0.154***	0.0797**	0.160***
	(0.0264)	(0.0312)	(0.0567)
Observations	57,682	79,876	7,586
R-squared	0.449	0.628	0.365
Panel B: Tariff based on initial industry			
Tariff	0.146***	0.0662**	0.179***
	(0.0198)	(0.0288)	(0.0486)
Observations	57,682	79,876	7,586
R-squared	0.450	0.628	0.365
Panel C: Tariff based on initial industry, with individual fixed effects			
Tariff	0.112***	0.0476**	0.0896**
	(0.0304)	(0.0238)	(0.0436)
Observations	57,682	79,876	7,586
R-squared	0.868	0.911	0.886

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The sample is based on workers age 20 to 64 as of the 2002 VHLSS that reported working in the 2002 and 2004 VHLSSs. The groupings into traded, all, and manufacturing are based on the initial industry of employment reported in the 2002 VHLSS. All regressions that do not include individual fixed effects include individual covariates (age, age squared, education levels, gender, ethnic minority status, urban indicator, and province fixed effects). All regressions include industry and year fixed effects.

Table 7: Industry Employment and Tariffs

Dependent variable: Share of industry employment in the indicated set of industries

	(1) Traded	(2) All	(3) Manufacturing
Panel A: Overall Employment (VHLSS Data)			
Tariff	-0.00464	-0.000263	0.00290
	(0.00555)	(0.00257)	(0.0321)
Observations	68	120	44
Within R-squared	0.023	0.000	0.001
Panel B: Enterprise Sector (Enterprise Survey Data)			
Tariff	-0.0265**	-0.0108**	-0.0257
	(0.0113)	(0.00533)	(0.0170)
Observations	66	110	44
Within R-squared	0.232	0.124	0.167

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The dependent variable is the share of workers and is calculated as the number of workers in industry j divided by the total number of workers in the respective group. The total number of workers includes workers in (i) traded industries for column (1), (ii) all industries for column (2), and (iii) traded manufacturing industries for column (3). In Panel A, the industry employment shares are based on the 2002 and 2004 VHLSSs and include workers between the ages of 20 and 64 inclusive. In Panel B, the industry employment shares are data from the 2000 and 2003 enterprises surveys. These employment estimates include all workers in enterprises at the end of 2000 and 2003 respectively. All regressions include year fixed effects and industry fixed effects, using the within transformation.

Table 8: Labor productivity gap per hour between the enterprise and household business sectors in manufacturing

	Manufacturing		Textiles and apparel		Ho Chi Minh City and Dong Nai	
	Revenue based	Wage based	Revenue based	Wage based	Revenue based	Wage based
	(1)	(2)	(3)	(4)	(5)	(6)
Labor productivity gap:						
Unadjusted	9.0	1.82	6.6	1.70	7.0	1.48
Adjusted by hours worked & human capital	6.0	1.24	4.7	1.28	5.5	1.15
+ measurement error in revenue and hours worked	3.7		2.9		3.4	
+ differences in output-labor elasticity	2.5		2.0		2.3	
Share of hours reallocated to enterprises due to the BTA	0.050	0.050	0.086	0.086	0.053	0.053
Initial share of hours in the household business sector	0.597	0.597	0.615	0.615	0.380	0.380
Annual growth (%)						
Adjusted by hours worked & human capital	3.5	0.5	5.8	1.0	2.7	0.3
+ measurement error in revenue and hours worked	2.8		4.3		2.2	
+ differences in output-labor elasticity	1.5		2.1		1.2	

Notes: The labor productivity gap reported in columns 1, 3, and 5 is based on the average revenue product of labor and subsequent adjustments. The average revenue product of labor is the ratio of revenue per worker in the enterprise sector to revenue per worker in the household business sector. The labor productivity gap reported in columns 2, 4, and 6 is based on the ratio of annual earnings per worker in the enterprise sector to annual earnings per worker in the household business sector, plus subsequent adjustments. The difference in output-labor elasticity allow the MRPL and ARPL gaps to differ. See section 6 and Appendix B for further details on the calculations and data sources.

Supplementary material for

Export markets and labor allocation in a low-income country

By Brian McCaig and Nina Pavcnik

(for online publication only)

Appendix Table A.1: Growth of Vietnamese exports and BTA tariff changes

Dependent variable: Change in ln exports

	(1)	(2)	(3)	(4)
Industries	Traded	Manufacturing	Traded	Manufacturing
Destination market	US	US	EU13	EU13
Panel A: Change in ln exports 2001 to 2004				
BTA tariff change	-5.677***	-4.331*	0.372	0.142
	(1.474)	(2.111)	(0.675)	(1.070)
Observations	24	19	24	19
R-squared	0.283	0.119	0.009	0.001
Panel B: Change in ln exports 1997 to 2000				
BTA tariff change	-0.808	0.181	0.362	0.823
	(1.896)	(1.722)	(0.599)	(0.904)
Observations	24	19	24	19
R-squared	0.011	0.001	0.011	0.035

Notes: Robust standard errors in parentheses. ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. We use data on imports from Vietnam as reported by the U.S. and EU13 (EU15 excluding Belgium and Luxembourgh for which data was not consistently available) in UNComtrade. We exclude industries for which imports were 0 for any of the years.

Appendix Table A.2: Descriptive statistics

	All		Pre BTA Round		Post BTA Round	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Self-employed	0.686	0.464	0.701	0.458	0.672	0.469
Worked in a household business	0.830	0.375	0.847	0.360	0.814	0.389
Indicator for urban	0.239	0.427	0.240	0.427	0.238	0.426
Age	37.8	11.1	37.4	11.0	38.3	11.1
Indicator for female	0.505	0.500	0.507	0.500	0.503	0.500
Indicator for ethnic minority	0.123	0.328	0.121	0.326	0.124	0.329
Indicator for completed primary education	0.288	0.453	0.297	0.457	0.280	0.449
Indicator for completed lower secondary education	0.300	0.458	0.292	0.455	0.307	0.461
Indicator for completed upper secondary education	0.199	0.399	0.185	0.388	0.212	0.409
Indicator for agriculture, forestry and aquaculture	0.542	0.498	0.561	0.496	0.524	0.499
Indicator for manufacturing	0.123	0.329	0.118	0.322	0.128	0.334
Indicator for services	0.327	0.469	0.313	0.464	0.341	0.474
Indicator for less than median distance from seaport	0.561	0.496	0.560	0.496	0.562	0.496
Hours per year (primary job)	1701	814	1730	792	1673	833
Indicator for more than one job					0.425	0.494
ln(hourly compensation)	1.368	0.726	1.234	0.791	1.494	0.634
Number of observations	248,795		152,388		96,407	

Notes: The sample consists of all workers from the 2002 and 2004 VHLSSs that worked and were 20 to 64 years of age inclusive at the time of the survey. The 2002 VHLSS is the pre BTA round and the 2004 VLHSS is the post BTA round. The number of observations for wages are lower: 46,309 and 29,758 in the 2002 and 2004 VHLSSs respectively. The total number of observations is slightly higher than in our regression results due to a small number of worker observations, 4, for which ethnicity data is missing and these observations are subsequently dropped from the regression analysis.

Appendix Table A.3: Differences in job characteristics between workers in household businesses and enterprises

	Traded		All		Manufacturing	
	Household businesses	Enterprises	Household businesses	Enterprises	Household businesses	Enterprises
Share self-employed	0.868	0.000	0.827	0.000	0.609	0.000
Mean annual hours worked	1,531	2,193	1,667	2,076	1,889	2,293
Mean highest grade completed	6.4	9.5	6.7	10.4	7.7	9.6
Share with more than one job	0.496	0.200	0.463	0.261	0.492	0.157
Mean annual total compensation	4,237	11,454	5,248	11,518	6,380	11,608
Share of wage workers that report receiving payments for …						
Holidays	0.150	0.846	0.188	0.861	0.296	0.863
Social insurance	0.001	0.034	0.002	0.037	0.003	0.032
Business trips	0.001	0.073	0.003	0.149	0.001	0.040
Other	0.117	0.585	0.156	0.605	0.168	0.606
Any non-wage/salary payment	0.215	0.895	0.273	0.914	0.362	0.913

Notes: The sample is all workers age 20 to 64 in the 2002 VHLSS, with the exception of the share of workers with more than one job, which is based on workers age 20 to 64 in the 2004 VHLSS. The values are estimated using sampling weights. Social insurance payments mean payments received by workers, for example, for workplace injuries, not necessarily social insurance coverage. Compensation is reported in 000s of Vietnamese dong in January 2002 prices.

Appendix Table A.4: Additional Results for Employment in Household Businesses and Tariffs
Dependent variable: Indicator for working in a household business

	(1) Traded	(2) All	(3) Manufacturing
Panel A: Worker controls are omitted			
Tariff	0.212***	0.126***	0.162***
	(0.0140)	(0.0342)	(0.0210)
Observations	176,546	248,793	27,072
R-squared	0.404	0.580	0.245
Panel B: Remove mining of uranium and thorium ores from sample			
Tariff	0.209***	0.128***	0.156***
	(0.0144)	(0.0323)	(0.0197)
Observations	176,538	248,785	27,072
R-squared	0.419	0.594	0.299
Panel C: Pre-existing trends included			
Tariff	0.210***	0.138***	0.160***
	(0.0141)	(0.0271)	(0.0177)
(Change in industry ln employment)	-0.0160**	-0.0266*	0.0100
X (2004 indicator)	(0.00736)	(0.0133)	(0.0161)
(Change in industry self-employment rate)	0.00196	0.0799	0.0124
X (2004 indicator)	(0.0554)	(0.0677)	(0.0595)
(Change in industry mean grade completed)	0.0117**	0.0115*	0.0163*
X (2004 indicator)	(0.00488)	(0.00601)	(0.00864)
Observations	176,455	248,355	26,981
R-squared	0.418	0.592	0.298

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The sample is restricted to workers between the ages of 20 and 64 inclusive at the time of the survey. Column (1) includes all traded industries, column (2) includes all industries, and column (3) includes all traded manufacturing industries. All regressions include a rural indicator, industry, province, and year fixed effects, and regressions in Panels B and C also worker characteristics (age, age squared, education level indicators, female indicator, and ethnic minority indicator). The pre-existing trends in Panel C are calculated between 1993 and 1998 using the respective Vietnam Living Standards Surveys. The decrease in the number of observations in Panel C is due to a few small industries with missing information for pre-existing trends.

Appendix Table A.5: Hours Worked in Household Businesses and Tariffs

Dependent variable: Share of total industry hours worked within the household business sector

	(1) Traded	(2) All	(3) Manufacturing
Tariff	0.202***	0.119***	0.157***
	(0.0186)	(0.0379)	(0.0302)
Observations	68	120	44
R-squared	0.565	0.350	0.814
Industry fixed effects?	Yes	Yes	Yes
Year fixed effect?	Yes	Yes	Yes

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. Column (1) includes all traded industries, column (2) includes all industries, and column (3) includes all traded manufacturing industries. Hours worked in the household business and enterprise sectors are estimated based on workers at 20 to 64 inclusive at the time of the 2002 and 2004 household surveys.

Appendix Table A.6: Industry Employment and Tariffs in Urban Areas

Dependent variable: Share of industry employment in the indicated set of industries

	(1) Traded	(2) All	(3) Manufacturing
Tariff	-0.00174	0.00163	0.00666
	(0.0136)	(0.00437)	(0.0421)
Observations	68	120	44
Within R-squared	0.001	0.006	0.003

Notes: Standard errors are clustered at industry level; ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The dependent variable is the share of workers and is calculated as the number of urban workers in industry j divided by the total number of urban workers in the respective group. The total number of urban workers includes workers in (i) traded industries for column (1), (ii) all industries for column (2), and (iii) traded manufacturing industries for column (3). The industry employment shares are based on urban workers between the ages of 20 and 64 inclusive, calculated from the VHLSSs. All regressions include year fixed effects and industry fixed effects, using the within transformation.

Appendix B: Supplemental Material for Section 6

This appendix provides a detailed description of the data and calculations used in section 6 to estimate the aggregate labor productivity change in manufacturing associated with the reallocation of labor from household businesses to enterprises in response to the BTA.

B.1 Labor productivity gap between the household business and enterprise sectors

B.1.1 Basic Calculations

We compute the average revenue product of labor based on revenue per worker from firm-level data that covers all registered firms in the enterprise sector (the Vietnamese Enterprise Survey), and on aggregate revenue and the total number of workers in the household business sector from the household business and labor modules of the VHLSS. Most of our analysis relies on the 2001 Vietnamese Enterprise survey and the 2002 VHLSS.[66]

We calculate average revenue product of labor in the household business sector by summing annual revenue from all businesses reported in the household business module and dividing by the total number of workers in household businesses as reported in the labor module of the 2002 VHLSS.[67,68] Our estimate of revenue per worker in the household business sector is likely an overestimate because the business module in the 2002 VHLSS did not distinguish between private enterprises (i.e., belonging in the enterprise sector) and household businesses. Its scope was *all* businesses run by households. This lowers the estimate of the gap in average labor productivity between the household business and enterprise sectors and thus leads to a cautious estimate of the gains from reallocation. See section B.1.5 for further discussion.

We calculate revenue per worker in the enterprise sector based on total annual revenue divided by total employment in each enterprise at year-end using the 2001 enterprise data.

[66] The VHLSS has has a 12-month recall and covers 2001/2002

[67] All revenue values are expressed in January 2002 Dong. For the VHLSS data, household business revenue is converted to January 2002 prices using monthly CPI data based on the month of interview. Information on the number of workers is based on employment in workers' primary job (i.e., employment based on their reported most time consuming job). See discussion of employment in secondary jobs in Section B.1.5. We rely on the labor force module for employment because the 2002 household business module did not collect information about the number of workers in each business.

[68] Both estimates are weighted using survey sampling weights.

Aggregate revenue per worker within the enterprise sector is the employment-weighted average of revenue per worker over all enterprises.

We use the labor force modules of the 2002 VHLSS to compute the wage gap between the enterprise and household business sectors. We compute the wage ratio using total earnings, which includes cash and in-kind wage/salary payments as well as other payments such as public holiday payments and social allowance payments, among wage workers in the two sectors.

These labor productivity gaps are reported in row 1 of Table 8 in the main text.

B.1.2 Existing evidence on the labor productivity gap

How do our estimates compare to those from the literature? Nationally representative data on household businesses (in other contexts also called informal firms or microenterprises) is rare, so we cannot compare our estimates to many previous studies. That said, our estimate is consistent with productivity gaps between informal and formal firms in other developing countries. We compare existing estimates to our unadjusted estimates in row one of Table 8 in the main text because the existing estimates do not adjust for worker heterogeneity across sectors. For example, Nataraj (2011) finds that output per worker is 12.4 times higher in formal firms than informal firms in India's manufacturing sector. La Porta and Shleifer (2014) report gaps in value added per worker between formal and informal firms using firm surveys from 25 developing countries. They rely on World Bank Enterprise Surveys, which are not nationally representative and often exclude firms without any hired workers (i.e., firms that the literature often refers to as own-account workers). Consequently, the informal firms are larger on average, employing four workers, than in our context where the average number of workers is about 2. A big advantage of our data is that it is based on a census of all registered firms and a survey of household businesses based on a nationally representative household survey. With these caveats about the World Bank Enterprise Survey data in mind, value added per worker in formal firms is 6.7 times higher than in informal firms in the median country in their sample. This is lower than in our data, but the range across countries is large, from 5.3 in Mali to 14.3 in Tanzania, the 25th and 75th percentile countries, respectively. Our estimate of the productivity gap of a factor of 9 is within the range of estimates from other countries, comparable to their estimates for Angola and Kenya.

B.1.3 Adjusting the labor productivity gap for worker heterogeneity

The productivity gap in row 1 of Table 8 could reflect worker heterogeneity across the two sectors. We use micro-survey data on individuals from the labor force and education modules of the VHLSS to adjust this productivity gap for heterogeneity in hours worked and human capital between the household business and the enterprise sectors.

To begin with, the estimates reported in row 1 of Table 8 are based on revenue per worker rather than revenue per hour worked. However, enterprise sector workers work significantly more hours in the year than household business workers, 2,267 versus 1,825.[69] The difference in hours will reduce the productivity gap. We adjust the productivity measures from row 1 so that they reflect revenue per hour worked in each sector.

Second, we also adjust the productivity gap for differences in human capital across sectors. We follow Gollin et al. (2014), who adjust for the differences in average years of education across the agriculture and non-agriculture sectors. Specifically, let average human capital within a sector be given by $\exp(rsch_s)$ where r is the rate of return to a year of schooling and sch_s is the average years of education in sector s. We use a rate of return of 10 percent as in Gollin et al. (2014), based on the observation in Banerjee and Duflo (2005) that returns to schooling are estimated to be around 10% in most countries and don't vary much between low- and high-income countries. Using this approach, human capital in the enterprise sector is 1.2 times higher than in the household business sector.

The productivity gap with both of these adjustments is reported in row 2 of Table 8 and discussed in the main text. These adjustments matter. As noted there, worker heterogeneity accounts for 37% of the original ARPL gap and for 70% of the original wage gap.

One potential objection to the above approach to controlling for worker heterogeneity is that it only controls for two dimensions of worker heterogeneity. Alternatively, one can estimate the wage gap for working in the enterprise sector with Mincerian regressions, while controlling for worker heterogeneity in other dimensions, including location, gender, age, ethnic minority status, and industry affiliation. Vollrath (2014) and Herrendorf and Schoellman (2015) apply this approach to the wage gap between non-agriculture and agriculture. We estimate the following regression:

$$\ln w_i = \alpha + \beta_e enterprise_i + \theta X_i + \varepsilon_i$$

[69] This is based on hours worked in primary job.

where w_i is the real hourly wage of worker i, $enterprise_i$ is an indicator variable for working in the enterprise sector, and X_i is a vector of worker characteristics, including education, location, gender, age, age squared, minority status, and industry affiliation. The wage gap for working in the enterprise sector (relative to the household business sector) is given by $\exp(\beta_e)$, where β_e is the coefficient on the indicator that a worker works in the enterprise sector.

We estimate the above Mincerian regression using data on manufacturing workers from the 2002 VHLSS. The estimates of the coefficient on the enterprise indicator β_e are reported in Appendix Table B.1. After controlling for additional dimensions of worker heterogeneity, the wage gap is a similar order of magnitude as the wage gap reported in Table 8. The estimates that are the most comparable to the wage ratio adjusted for human capital in row 2 of Table 8 are in column 2.[70] The estimates in the table are of a similar order of magnitude as the estimates based on the wage ratio adjusted for worker characteristics.[71] For example, the most conservative estimate of the wage difference, reported in column 6, which controls for education, location, demographics, and industry affiliation, suggests that workers in the enterprise sector earn 25% more per hour than observationally equivalent workers in the household business sector. Thus, estimates of the wage gap based on the Mincerian regressions, which simultaneously control for several dimensions of worker heterogeneity, yield similar findings as the approach taken in the text.

We also use the individual panel data to estimate the impact on wages while controlling for unobserved worker heterogeneity, which might be correlated with worker wages and sector of employment. The results are reported in Appendix Table B.2. Our sample is all individuals that worked for wages in manufacturing in *both* 2001/02 and 2003/04. Hence, the number of workers included is lower than in Appendix Table B.1. The estimated coefficients on an indicator for working in the enterprise sector from a regression that controls for unobserved worker heterogeneity with individual fixed effects are reported in columns 1 and 2. The wage gap persists, albeit it is smaller in magnitude. The estimates of the coefficient β_e suggest that wage workers that move between these

[70] Comparisons between estimates in column 2, which control for heterogeneity in education with years of education, and column 3, which control for heterogeneity in education with education indicators, suggest that these two approaches yield similar coefficient on the enterprise sector indicator. We therefore control for heterogeneity in education with education indicators, as we did in Section 5 of the paper.

[71] The difference in the estimates between the wage gap and Mincerian-based estimates could be attributed to additional controls for observable worker characteristics, differences in functional form assumption, and the fact that the wage bill ratio is weighed by hours worked, while the Mincerian regression weights each hourly wage observation equally.

two sectors earn about 9% more per hour when working in the enterprise sector. As the sample has changed relative to Appendix Table B.1, we also estimate β_e using the cross-sectional specification that uses the same set of worker covariates as column 6 in Appendix Table B.1 with the panel sample. The results in column 3 are based on both 2002 and 2004 data, while column 4 uses only 2002 data. The estimates of β_e are very similar to those using the cross section in column 6 of Appendix Table B.1 (0.202 and 0.191 versus 0.221). Hence, the change in sample is not the primary reason for the lower coefficient on working in the enterprise sector reported in columns 1 and 2. Instead, controlling for unobserved individual heterogeneity is the important driver.

The above analysis excludes self-employed individuals, as they do not work for wages. Hence, as an additional check relative to the wage regressions, we also include the self-employed in the above analysis by focusing on worker hourly income as a dependent variable. For wage workers, this is wage earnings as above. For the self-employed, we use self-reported profits. Specifically, we focus on manufacturing household businesses run by the manager as their primary job. Additionally, we restrict the sample of businesses to those for which no other household member reports being self-employed in the same industry. This allows us to assign the reported profits to the manager without having to make assumptions on how to assign profits across multiple household members. In this analysis, we are treating the profits from the business as the manager's wage earnings.

These results are reported in Appendix Table B.3, which follows the same specifications as Appendix Table B.1. Interestingly, while the findings are in general similar to the findings with wages, the earnings gap is smaller in magnitude, at 11 percentage points, relative to the differences in wages only. This suggests that the managers of household businesses earn higher profits than observationally equivalent wage workers.

In summary, the above discussion provides the details of how we adjust the labor productivity gaps for worker heterogeneity. In the main text, we use these insights to emphasize that worker heterogeneity accounts for an important component of the ARPL and wage gaps across sectors, 37 % and 70%, respectively. For the wage gap, worker heterogeneity plays an important role when we estimate the gap with differences in mean earnings between sectors, analogous to the ARPL gap, and when we estimate the gap using Mincerian wage or earnings regressions. Overall, this analysis highlights the importance of relying on information from micro-survey data to account for worker heterogeneity.

B.1.4 Measurement concerns

We address several concerns about measurement error. To begin with, one may be concerned about the measurement error associated with combining two different data sources.[72] For example, the surveys might measure revenue or earnings differently and this affects the estimates of the productivity gap. We perform several checks to address this potential concern.

First, we compare annual earnings per manufacturing worker (i.e., the numerator of the wage gap) in the enterprise sector across the two data sources, the 2002 VHLSS and the 2001 enterprise data. Mean annual earnings per worker were 11.6 million VND in the 2002 VHLSS as compared to 11.8 million VND in the 2001 enterprise data. The similarity of these estimates suggest that workers in the 2002 VHLSS are reporting labor earnings consistent with the reports of labor expenses from firms in the 2001 enterprise data. Thus, the wage gap estimates between the enterprise and household business sector are very similar if we use wage data for enterprises based on VHLSS or the enterprise survey.

Second, the business module of a later VHLSS survey, the 2006 VHLSS, collected information on whether the business owned by a household was a household business or registered as an enterprise. We use this data to estimate the ARPL gap between household businesses and private enterprises with one survey. 1.5 percent of manufacturing businesses in the 2006 VHLSS were private enterprises, emphasizing the prevalence of household businesses in manufacturing. Note that this calculation does not capture the productivity gap between household businesses and other types of firms in the enterprise sector. Specifically, it excludes state-owned enterprises, foreign-invested firms, and collective firms. We estimate the ARPL gap between private enterprises and household businesses using the business module of the 2006 VHLSS.[73] These results are reported in column 1 of Appendix Table B.4. ARPL is 3.6 times higher among private, domestic enterprises than household businesses. The ratio drops to 2.8 after we adjust it for worker heterogeneity. The gap is slightly lower than in Table 8 since foreign-invested, state-owned, and collective enterprises are not included in the estimate of average productivity in the enterprise sector. This gap is computed with data from 4 years after the 2002 VHLSS baseline year. For comparison purposes, we estimate the ARPL gap between the entire enterprise sector and

[72] See Nataraj (2011) and Hsieh and Olken (2014) for such an approach using Indian data and Ulyssea (2017) in Brazil.
[73] The 2006 VHLSS has a 12-month recall and was conducted from May through November.

household business sector for 2005/06, using the same procedure from section B.1.1 and the same two data sources for 2005/06 as those used for 2002 gaps reported in Table 8. These estimates are reported in columns 2 through 4 of Appendix Table B.4. For example, the estimates of the ARPL gap for manufacturing as a whole in column 2 suggest that the ARPL gap between sectors is 5.0, 3.5 adjusted for worker heterogeneity. As expected, they are higher than the estimates that rely on private domestic enterprises alone. In addition, these estimates suggest that the ARPL gap has partially closed between 2001/02 and 2005/06.

Overall, the above discussion suggests that the productivity gap between workers in the enterprise and household business sectors is not likely driven entirely by differences in how firms report revenues in the VHLSS and the enterprise survey.

An alternative measurement concern is that very few microenterprise operators keep formal accounts and thus measurement error is more likely to affect our estimate of ARPL in the household business sector than in the enterprise sector. We discuss the potential measurement error in revenue and labor supply and consider adjustments to the ARPL gap to take these concerns into account. The results are reported in column 1 of Appendix Table B.5. The first row of the table reports the ARPL gap adjusted for worker heterogeneity that was discussed in Appendix B.1.3 and that is also reported in row 2, column 1 of Table 8. This estimate serves as the upper bound on the productivity gap.

First, de Mel, McKenzie and Woodruff (2009) provide experimental evidence on possible measurement error in revenues of microenterprises from Sri Lanka, suggesting that reported revenue in these businesses may underestimate true revenue by as much as 30 percent. Their sample covers microenterprises without any paid employees, in both retail and manufacturing. When we adjust the reported revenue in the household business sector by this factor, the productivity gap falls from 6.0 to 4.2 (column1, row 2 of Table B.5).

Second, people working in the household business sector might overstate effective hours worked. For example, a shopkeeper might be watching her children while tending to the business, but reports total hours worked in the shop in the survey. We are not aware of a study that examines potential mismeasurement of labor supply to microenterprises. However, the data from Ghana generated by Fafchamps, McKenzie, Quinn and Woodruff (2014), which covers a sample of microenterprises without paid employees, asked about the number of hours worked last week as

well as the number of hours worked at full effort.[74] Interestingly, the difference in hours worked versus full effort hours is not very large. Across all rounds of the surveys, microenterprise owners reported working at full effort 89 percent of the time. This suggests that reports of hours worked may slightly overestimate true labor input. When we adjust the productivity gap for the measurement error in revenue and hours worked, the ARPL gap is 3.7 (column 1, row 3 of Table B.5).

B.1.5 Measurement issues specific to the 2002 VHLSS

The above discussion focuses on measurement issues related to informal businesses in any data. In this section, we use the 2004 and 2006 VHLSSs data to assess two measurement issues specific to the 2002 VHLSS data used in the current study.

First, our estimate of aggregate revenue in the household business sector is likely overestimated due to the inclusion of some private enterprises in the estimate. The 2002 VHLSS asked each household whether they ran a business, but it did not distinguish between household businesses and private enterprises. Consequently, revenue from private enterprises is also included in our estimate of aggregate revenue in the household business sector. However, the labor module clearly distinguishes between working in a household business versus a private enterprise. Hence, our estimate of ARPL in the household business sector is an overestimate, as it is based on the revenue from household businesses and private enterprises. This underestimates the ARPL gap between the enterprise and household business sectors.

The 2006 VHLSS distinguished between the two types of private businesses, household businesses and private enterprises, and we use this survey to get an estimate of how much we might be overstating aggregate revenue from the household business sector by including private enterprises. The estimate based on the 2006 data likely overstates the contribution of private enterprises to the 2002 household business revenue because the relative share of the formal sector has been growing over time. As a result, the adjustments below should be viewed as providing a lower-bound estimate of the ARPL. Only 1.5 percent of manufacturing businesses in the 2006 VHLSS were private enterprises. However, they tend to be bigger. In the 2006 VHLSS, among

[74] This question was asked because many microenterprise owners are simultaneously engaging in non-microenterprise activities, such as childcare or household work, while operating their business. We thank Christopher Woodruff for making us aware of this data.

manufacturing businesses, 23% of aggregate revenue is from private enterprises. Under the assumption that this share has not changed over time, we adjust the ARPL ratio for this measurement error by subtracting the same proportion of revenue from the 2002 VHLSS based estimates. The ARPL gap rises to 4.9 (column 1, row 4 of Table B.5).

In addition, our estimate of hours worked in the household business sector omits hours worked as a secondary job because the 2002 VHLSS did not collect detailed data on secondary jobs, defined as the second most time consuming job during the past 12 months. This underestimates labor supplied to the household business sector. We assess the potential measurement error with the 2004 and 2006 VHLSSs, which asked detailed questions about both the primary and secondary jobs of workers. In the enterprise sector, only 1.2% and 1.0% of total hours reported in the 2004 and 2006 VHLSSs were worked in secondary jobs. However, 13.9% and 13.8% of total hours in the household business sector in the 2004 and 2006 VHLSSs were worked as a secondary job. Adjusting the 2002 estimate of the ARPL ratio to include the missing hours in the household business sector due to secondary jobs increases the estimate to 5.6 (column1, row 5 of Table B.5).

In summary, while measurement error is certainly present in our data, attempts to adjust for plausible measurement error do not eliminate the ARPL gap across the household business and enterprise sectors within manufacturing. In the main text we focus on the most conservative estimate of the productivity gap from column 1 of Table B.5, namely 3.7.

B.1.6 Differences in the output elasticity of labor

As equation (4) in section 6 makes clear, the ARPL gap may overestimate the MRPL gap if the household business sector has a higher output elasticity of labor, because $\frac{MRPL_e}{MRPL_h} = \frac{(1-\alpha_e)}{(1-\alpha_h)}\frac{ARPL_e}{ARPL_h}$. Estimating production functions is beyond the scope of this paper. Hence, we use existing values from the literature. There are not many studies from which to draw. Fernandez and Meza (2015) calibrate a model of informal and formal firms using output elasticities of labor of 0.8 and 0.65, respectively. Restrepo-Echavarria (2014) assumes output-labor elasticities of 1 and 0.68 in the informal and formal sectors, respectively. [75] We choose the largest elasticity

[75] Restrepo-Echavarria (2014) reports value added-labor elasticity. Adjusting for factor share of materials based on Nataraj (2011) yields similar results because the factor share of materials is very similar across the two sectors.

ratio of about 1.5 (approximately 1/0.68) and apply it to the ARPL gaps reported in column 1 of Table B.5. The obtained MRPL ratios are reported in column 3 of Table B.5.

Let us first focus on the ARPL gap that adjusts for worker heterogeneity and hours work differences across the two sectors and is reported in row 1. Adjusting this ratio with the elasticity ratio suggests an MRPL ratio of about 4 (row 1, column 3 of Table B.5). MRPL ratios associated with alternative estimates of labor productivity gaps in column 1 of Table B.5 discussed in section B.1.5 are reported in the remaining rows in column 3. The most conservative MRPL ratio of 2.5 is the one reported in row 3 and based on the productivity gap of 3.7 that adjusts for worker heterogeneity, and the measurement error in revenue and hours worked. Note that we used a factor of 1.5, a very large difference in elasticities, to provide a conservative estimate of the possible MRPL gap. In the main text, we therefore discuss this most conservative estimate as the lower bound for the labor productivity gap, adjusted for measurement error and differences in output-labor elasticities across sectors.

B.2 Estimating the share of workers reallocated

The estimate of the aggregate labor productivity gain in section 6 of the main text requires an estimate of the share of manufacturing workers reallocated from the household business to the enterprise sector due to the BTA, S^{BTA}.

Within each industry j, we estimate the share of workers reallocated as

$$\hat{\beta} \times \Delta tariff_j$$

where $\hat{\beta}$ is estimated based on equation (3) and $\Delta tariff_j$ is the change in the U.S. tariff on Vietnamese exports in industry j due to the BTA. We then sum over manufacturing industries, weighting by the industry's share of overall manufacturing employment:

$$\sum_j \left(\hat{\beta} \times \Delta tariff_j \right) s_j$$

where s_j is the share of manufacturing workers in industry j. To be consistent with how we calculate the number of workers for the ARPL gaps, we calculate total employment in each industry as the sum of household business sector employment, estimated from the 2002 VHLSS, and

enterprise sector employment, derived from employment at the end of 2001 from the enterprise data.[76]

Our preferred estimate for $\hat{\beta}$ is 0.156 (see Panel A, column 3 of Table 3). With this value, we estimate that 4.9 percent of manufacturing workers were reallocated out of the household business sector to the enterprise sector. The regression estimates reported in Table 3 focus on the reallocation of workers between sectors, whereas section 6 focuses on labor productivity per hour worked, not per worker. In our case, this distinction in the reallocation of labor, whether workers or hours, turns out to be inconsequential as the BTA induced essentially an identical change in the share of hours worked in the household business sector (see the coefficient of 0.157 in column 3 of Table A.5 for share of hours worked as compared to our benchmark estimate of 0.156 from Table 3). The estimates of reallocated hours worked are reported in Table 8 in the main text. They are used in all subsequent calculations discussed in section B.3.

B.3 Aggregate Labor Productivity Gain

We use the formula in section 6 of the main text, equation (5), to compute the gain in labor productivity based on various labor productivity gaps. Recall that these labor productivity gaps are reported in Appendix Table B.4 and column 1 and 3 of Table B.5. Those tables also report the associated annual aggregate labor productivity gaps.

Here we focus our discussion on the estimates of the aggregate gains for the labor productivity gaps reported in Appendix Table B.5. Recall that in column 1, row 1, we present an ARPL gap of 6.0, which adjusts for worker heterogeneity. Subsequent rows sequentially adjust the ARPL gap for each measurement error issue discussed previously. The ARPL gap falls as we consider possible reporting error in household business revenue and labor input and then increases as we remove possible revenue from private enterprises being included as household business revenue and add additional labor inputs from secondary jobs. These movements in the ARPL gap are reflected in the estimates of aggregate labor productivity gains. Our preferred labor productivity gap estimate of 3.7 in row 3 suggests that aggregate labor productivity per hour worked increased by 2.8% annually over the period of the study. This is the estimate also reported in Table 8 and discussed in section 6 of the main text. We view the 3.5% estimate in row 1, column 2 (which is

[76] If we estimate industry employment solely from the 2002 VHLSS we arrive at a similar estimate of the share of manufacturing workers reallocated.

based on the labor productivity gap that only adjusts for worker heterogeneity) as the upper bound on the aggregate labor productivity gains. Overall, the estimated productivity gains range from 2.8 to 3.5% per year.[77]

Column 4 of Appendix Table B.5 presents estimates of the annual growth in labor productivity that take into account potential measurement issues and additionally allow for differences in the output-labor elasticity across the household business and enterprise sectors. These are based on the MRPL gaps reported in column 3. The growth estimates range from 1.5 to 2.1% annually. Our most conservative estimate of the aggregate productivity gain is 1.5% (row 3, column 4). We discuss this gap in the main text and Table 8.

In Appendix Table B.6, we replicate the analysis of Table 8 except that labor productivity is based on per worker instead of per hour worked. We do so because micro data on hours worked is often not available in firm-level datasets and thus labor inputs are typically measured on a per worker basis. We focus our discussion on column 1, which is based on all of manufacturing. The unadjusted labor productivity gap is 9, as in Table 8. Subsequently, we adjust for differences in human capital, but not in hours worked and estimate a labor productivity gap of 7.5. Additional adjustments for measurement error in revenue and labor inputs for household businesses as well as the difference in the output-labor elasticity across sector are the same as in Table 8 for labor productivity per hour. At the bottom of column 1 in Table B.6, we report the aggregate labor productivity estimates due to the BTA associated with the reported labor productivity gaps. These range between 2.0 and 4.1% annually, with an estimate of 3.4% as our preferred estimate of the gains in aggregate labor productivity per worker. This preferred estimate parallels that for aggregate labor productivity per hour in that the associated labor productivity gap in both cases is based on adjustments for measurement error in revenue and labor inputs for household businesses, but assumes the same output-labor elasticity across sectors.

We can also use equation (5) to estimate the growth in wages due to the BTA-induced reallocation of workers towards formal firms within manufacturing. To do so, we simply use the labor productivity gap based on the wage gap and use initial wage levels in each sector instead of labor productivity levels. The wage gap adjusted for worker heterogeneity and annual hours worked

[77] In Appendix Table B.4, we also report growth in aggregate labor productivity based on labor productivity gap using data from four years later. Although the labor productivity gap is lower, 3.5 versus 6.0 after adjusting for differences in human capital and hours worked, the associated growth in aggregate labor productivity is 2.7% per year.

is reported in row 2, column 2 of Table. The associated growth in wages per hour worked is 0.5% annually. In Appendix Table B.6, we report the wage gap based on annual earnings (i.e., not adjusting for differences in hours worked between the informal and formal sectors), but adjusted for differences in human capital (row 2, column 2). The associated gain in annual wages is 0.9% per year.

Appendix Table B.1: Differences in hourly wages between enterprise and household business sector workers in manufacturing

Dependent variable: ln real hourly wage

	(1)	(2)	(3)	(4)	(5)	(6)
Enterprise sector indicator	0.366***	0.296***	0.291***	0.160***	0.227***	0.221***
	(0.0129)	(0.0133)	(0.0133)	(0.0131)	(0.0127)	(0.0136)
Observations	9,416	9,416	9,416	9,416	9,416	9,416
R-squared	0.079	0.101	0.104	0.280	0.356	0.372
Location?	No	No	No	Yes	Yes	Yes
Demographics?	No	No	No	No	Yes	Yes
Years of education?	No	Yes	No	No	No	No
Education categories?	No	No	Yes	Yes	Yes	Yes
Industry FEs?	No	No	No	No	No	Yes

Notes: ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The sample is all wage workers in manufacturing age 20 to 64 in the 2002 VHLSS. Location controls include a rural identifier and province fixed effects. Demographic controls include age, age squared, a female indicator, and an ethnic minority indicator. Education controls include indicators for completed primary, completed lower secondary, and completed upper secondary with did not complete primary as the excluded category. Industry controls include industry fixed effects.

Appendix Table B.2: Differences in hourly wages between enterprise and household business sector workers in manufacturing

Dependent variable: ln real hourly wage

	(1)	(2)	(3)	(4)
Enterprise sector indicator	0.0874*	0.0892*	0.202***	0.191***
	(0.0459)	(0.0463)	(0.0234)	(0.0354)
Observations	2,680	2,680	2,680	1,340
R-squared	0.052	0.077	0.435	0.445
Number of individuals	1,340	1,340	1,340	1,340
Location?	No	No	Yes	Yes
Demographics?	No	No	Yes	Yes
Years of education?	No	No	No	No
Education categories?	No	No	Yes	Yes
Industry?	No	Yes	Yes	Yes
Year FEs?	Yes	Yes	Yes	No
Individual FEs?	Yes	Yes	No	No

Notes: ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The sample is a panel of workers age 20 to 64 in the 2002 VHLSS that worked in manufacturing in 2002 and 2004 for a wage. In column (4) the analysis is restricted to 2002.

Appendix Table B.3: Differences in hourly earnings between enterprise and household business sector workers in manufacturing
Dependent variable: ln real hourly earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Enterprise sector indicator	0.304***	0.235***	0.225***	0.0713***	0.126***	0.112***	0.139***	0.338***	0.261***
	(0.0126)	(0.0131)	(0.0131)	(0.0130)	(0.0123)	(0.0133)	(0.0126)	(0.0120)	(0.0140)
Observations	12,943	12,943	12,943	12,943	12,943	12,943	12,943	12,943	12,943
R-squared	0.039	0.059	0.062	0.219	0.288	0.302	0.195	0.117	0.090
Location?	No	No	No	Yes	Yes	Yes	Yes	No	No
Demographics?	No	No	No	No	Yes	Yes	No	Yes	No
Years of education?	No	Yes	No	No	No	No	No	No	No
Education categories?	No	No	Yes	Yes	Yes	Yes	No	No	No
Industry FEs?	No	No	No	No	No	Yes	No	No	Yes

Notes: ***, **, and * denotes significance at 1, 5, and 10 percent level, respectively. The sample is all wage workers and managers of one-worker household businesses in manufacturing age 20 to 64 in the 2002 VHLSS. Location controls include a rural identifier and province fixed effects. Demographic controls include age, age squared, a female indicator, and an ethnic minority indicator. Education controls include indicators for completed primary, completed lower secondary, and completed upper secondary with did not complete primary as the excluded category. Industry controls include industry fixed effects.

Appendix Table B.4: Productivity gap between the enterprise and household business sectors in manufacturing in 2005

	Average revenue product of labor			
	Manuf. (2006 VHLSS) (1)	Manuf. (2)	Textiles and apparel (3)	Ho Chi Minh City and Dong Nai (4)
Productivity gap	3.6	5.0	5.0	3.3
Productivity gap adjusted by hours worked & human capital	2.8	3.5	3.9	2.6
Share of hours reallocated to enterprises due to the BTA	0.050	0.050	0.086	0.053
Initial share of hours in the household business sector	0.597	0.597	0.615	0.380
Annual growth in revenue per hour worked (%)	2.4	2.7	5.3	1.8

Notes: The productivity gap for the average revenue product of labor is the ratio of revenue per worker in the enterprise sector to revenue per worker in the household business sector. In column 1, we report the ratio for private enterprises versus household businesses as calculated from the business module in the 2006 VHLSS. In columns 2 through 4, we report the ratio for all enterprises, as calculated using the 2005 enterprise data, versus household businesses from the 2006 VHLSS. See section 6 and Appendix B for further details on the calculations and data sources.

Appendix Table B.5: Adjusting the average revenue product of labor gap

Adjusted for:	Measurement error		Measurement error and output-labor elasticity	
	ARPL gap	Annual growth rate (%)	MRPL gap	Annual growth rate (%)
Adjusted for hours and human capital	6.0	3.5	4.0	2.1
+ Adjusted for underreporting of microenterprise revenue by 30%	4.2	3.0	2.8	1.7
+ Adjusted for hours worked at full effort	3.7	2.8	2.5	1.5
+ Adjusted for revenue from included private enterprises	4.9	3.2	3.2	1.9
+ Adjusted for hours worked in a secondary job	5.6	3.4	3.8	2.1

Notes: The table reports the ARPL gap between the enterprise and household business sectors adjusted for measurement error and differences in output-labor elasticities across sectors. Each row is based on the adjustment described in the row and all adjustments in previous rows. It also reports the annual growth rate of hourly labor productivity associated with each adjusted ARPL gap.

Appendix Table B.6: Productivity gap per worker between the enterprise and household business sectors in manufacturing

	Manufacturing		Textiles and apparel		Ho Chi Minh City and Dong Nai	
	Revenue based	Wage based	Revenue based	Wage based	Revenue based	Wage based
	(1)	(2)	(3)	(4)	(5)	(6)
Productivity gap						
Unadjusted	9.0	1.82	6.6	1.70	7.0	1.48
Adjusted by human capital	7.5	1.47	5.8	1.51	5.9	1.20
+ measurement error in revenue and hours worked	4.6		3.6		3.7	
+ differences in output-labor elasticity	3.1		2.4		2.4	
Share of workers reallocated to enterprises due to the BTA	0.049	0.049	0.085	0.085	0.053	0.053
Initial share of workers in the household business sector	0.648	0.648	0.664	0.664	0.395	0.395
Annual growth (%)						
Adjusted by human capital	4.1	0.9	6.9	1.7	2.7	0.4
+ measurement error in revenue and hours worked	3.4		5.4		2.3	
+ differences in output-labor elasticity	2.0		2.9		1.2	

Notes: The labor productivity gap reported in columns 1, 3, and 5 is based on the average revenue product of labor and subsequent adjustments. The average revenue product of labor is the ratio of revenue per worker in the enterprise sector to revenue per worker in the household business sector. The labor productivity gap reported in columns 2, 4, and 6 is based on the ratio of annual earnings per worker in the enterprise sector to annual earnings per worker in the household business sector, plus subsequent adjustments. The difference in output-labor elasticity allow the MRPL and ARPL gaps to differ. See section 6 and Appendix B for further details on the calculations and data sources.