UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2018/114

Market constraints, misallocation, and productivity in Vietnam agriculture

Loren Brandt,[1] Stephen Ayerst,[1] and Diego Restuccia[2]

September 2018

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: We examine important changes in agriculture in Vietnam in the context of ongoing structural changes in the economy. We use a household-level panel dataset and a quantitative framework to document the extent and consequences of factor misallocation in agriculture during the period between 2006 and 2016. Despite rapid growth in agricultural productivity and a reallocation of factor inputs to more productive farmers, we find that misallocation across farmers remains high and increased during the period. Reallocation of factor inputs has not been strong enough to accommodate substantial changes in farm productivity over time. Our analysis also reveals important differences between the north and south regions.

Keywords: agriculture, misallocation, Vietnam, productivity, regions
JEL classification: O11, O14, O4, E02, Q1

Acknowledgements: For useful comments and suggestions we are grateful to workshop participants in Helsinki. Diego Restuccia gratefully acknowledges the support from the Canada Research Chairs programme. Loren Brandt thanks the support from the Noranda Chair for International Trade and Economics. Brandt and Restuccia gratefully acknowledge the support from the Social Sciences and Humanities Research Council of Canada.

[1] University of Toronto, Toronto, Canada; [2] University of Toronto, Toronto, Canada, and NBER, Cambridge, MA, USA; corresponding author: Loren Brandt, email: brandt@chass.utoronto.ca.

This study has been prepared within the UNU-WIDER project on 'Structural transformation and inclusive growth in Vietnam'.

Copyright © UNU-WIDER 2018

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-556-5 https://doi.org/10.35188/UNU-WIDER/2018/556-5

Typescript prepared by Gary Smith.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

1 Introduction

Over the period between 2006 and 2016, Viet Nam enjoyed annual growth in gross domestic product (GDP) of 6 per cent and in labour productivity of 3.7 per cent. This growth was accompanied by higher growth in industry and services and a substantial shift of labour out of agriculture. In this paper we examine important changes in agriculture in Viet Nam in the context of this development in the economy. A rapidly expanding literature considers the role of agriculture in the process of structural transformation, and the consequences of resource misallocation and low productivity growth in the sector for how rapidly it proceeds.

Several basic questions motivate our analysis. How well does agriculture in Viet Nam perform over this 10-year period? How important is productivity growth, and the contributions of the intensive versus extensive margins? In the context of a rapid increase in off-farm demand for labour, how successful are factor markets in land, labour, capital, and intermediate inputs in facilitating reallocation of farm inputs to the most productive of farmers and uses? Is misallocation a serious problem in the farm sector? If so, how is it tied to local institutions? Can its behaviour be linked to productivity changes? Finally, are there important regional differences in how these processes are unfolding?

To address these questions, we draw on biennial household data from the VARHS (Vietnam Access to Resources Household Survey) that covers most of Viet Nam and provides a balanced panel of 2,087 households for the period between 2006 and 2016. We focus on the cropping sector, which consistently represents upwards of two-thirds of income in the agricultural sector, with animal husbandry, aquaculture, and forestry making up the rest. At the beginning of this period, 1,823 out of the 2,087 households were engaged in farming; by 2016, the number had fallen to 1,581, implying an annual exodus from agriculture of 1 per cent. More rapid reductions in labour supply are observed in terms of either the number of individuals working in agriculture, or the total number of days supplied to agriculture.

Our estimates suggest relatively rapid growth in agricultural output over the balanced panel of households, averaging more than 4 per cent per annum. Moreover, all of this growth arises from productivity improvements as increases in farm use of intermediate inputs such as fertilizer are more than offset by a reduction in land, labour, and capital. These productivity gains are accompanied by a shift to higher-valued crops, notably perennials, expanding average farm plot size, and a slight shift to larger farms.

Despite these gains, misallocation in the cropping sector is high and likely rising. We estimate the increases in output that could be obtained through allocating farm inputs to their first-best use, and find that this is rising over time. Although aggregate productivity in agriculture is improving, widening dispersion of farm productivity among households and frictions in input markets are a source of rising misallocation. Intuitively, the reallocation of resources in the farm sector is not keeping up with the rapid changes we observe in productivity at the household level.

Significant differences also emerge between the north and south regions of Viet Nam. In fact, all of the growth in farm output in our sample is coming from the south; there is none in the north. Differences also emerge with respect to productivity growth, which is almost two times higher in the south than in the north. Consistent with these estimates, we find much larger increases in our measures of resource misallocation in the north compared to the south.

Our analysis points to continued institutional constraints at the local and regional levels that impede the flow of resources to the most productive farmers. These constraints have potentially important long-run implications for productivity growth as they may limit the incentives for households to undertake productivity-enhancing investment in land, new crops, and technology.

Our paper is related to the broad literature on misallocation (e.g., Hsieh and Klenow 2009; Restuccia and Rogerson 2008).[1] More specifically, our paper relates to the recent literature emphasizing the importance of frictions in land and labour markets as sources of misallocation in the agricultural sector (e.g., Adamopoulos and Restuccia 2014; Adamopoulos et al. 2017; Ngai et al. 2017; Restuccia and Santaeulalia-Llopis 2017). Our paper also relates to an expanding literature examining the impact of policy on Viet Nam's agriculture sector (Tarp 2017).

The paper proceeds as follows. Section 2 provides broad background and economic context for Viet Nam and important details of the institutional environment, with a particular focus on constraints to factor allocation across households and sectors. In Section 3, we describe the household micro panel data from Viet Nam that we use and explain the construction of key variables. Section 4 examines changes in our household panel data in aggregate farm output and input use, and we carry out a growth accounting exercise to identify the role of the intensive and extensive margins in the changes in output we observe. In Section 5, we describe the main framework for measuring misallocation. Section 6 presents the main results on factor misallocation over this period, and characterizes key differences between regions in Viet Nam. In Section 7 we attempt to provide a link between the likely institutional constraints on household choices and misallocation, discussing the potential impediments to resource reallocation and growth. Section 8 provides robustness analysis with respect to variations in land quality as an alternative determinant of productivity differences across households. We conclude in Section 9.

2 Background: agriculture in Viet Nam

Our analysis begins in 2006, nearly two decades after the onset of economic reform in Viet Nam in the late 1980s. A brief examination of earlier rural reforms provides valuable institutional context.

At the heart of these efforts in the countryside were the decentralization of farming to the household and liberalization of input and output markets. At the beginning of the reforms, more than 90 per cent of all households in Viet Nam resided in the countryside. In the late 1980s, production rights to land reverted to households, and over time expanded to include rights to transfer, exchange, lease, inherit, and mortgage. Titling of land began in 1994 with the passing of the 1993 Land Law, and by 1997 Land Use Certificates (LUCs) had been issued to approximately one-half of all cultivated land (Benjamin and Brandt 2004). LUCs provided secure tenure for 20 years in the case of annual land, and 50 years for perennial.[2] By 2004, coverage extended to three-quarters of all cultivated land (Brandt et al. 2006), but over the next 10 years failed to increase (Markussen 2017). Important regional differences also persist. Do and Iyer (2008) and Newman et al. (2015) link land titling to rising investment in land.

Property rights reforms were accompanied by liberalization of product markets, especially for rice, and input markets such as those for fertilizer (Benjamin and Brandt 2004). Restrictions on the volume of rice exports were relaxed, as were internal product market barriers. Similarly, restrictions on fertilizer imports were removed. Prices came to be largely market-determined, and geographic mobility barriers were also relaxed. Estimates from the 2009 Population Census show the migration between provinces, much of it from the countryside to the cities, increasing from 1.3 million in 1989, to 2.0 million in 1999, and 3.4 million in 2009, or 4.3 per cent of the population (Narciso 2017). A key driver of migration decisions was growing opportunities in the secondary and tertiary sectors resulting from increases in inward foreign direct investment (FDI), expanded external market access tied to the United States-Viet

[1] See Restuccia and Rogerson (2017) for a general discussion of this literature.

[2] The 2013 Land Law conferred use rights for 50 years for all types of farmland.

Nam Bilateral Trade Agreement in 2000 and Viet Nam's accession to the World Trade Organization (WTO) in 2007, and state-owned enterprise (SOE) reform (McCaig and Pavcnik 2013).

These reforms contributed to impressive growth in agriculture, but a review of the recent literature suggests that constraints on household decision-making and resource allocation remain. The sector continues to be handicapped by a combination of government-imposed restrictions on farm size and the uses of agricultural land, and extensive land-use planning (World Bank 2016). Restrictions on crop choice persist, largely related to rice production and national food security (Markussen 2017). State involvement in agricultural value chains is also heavy. Furthermore, access to water for agriculture remains controlled by the government through irrigation SOEs. Low water productivity in the sector has been linked to limited pricing of irrigation water and monitoring of water use (World Bank 2016). Finally, land markets, especially those related to the buying and selling of land, remain thin as a result of high transaction costs. Households also face risks of land expropriation, with these risks negatively related to a household's informal ties to local officials and cadres (Markussen and Tarp 2014).

Tables 1 and 2 provide a breakdown of GDP and employment by sector between 2006 and 2016. By 2006, agriculture's share of GDP had fallen to 20 per cent. Its share of employment also declined, but much less so, and in 2006 more than half of the labour force, or nearly 24 million individuals, were still working in agriculture. Over the next 10 years, Viet Nam enjoyed aggregate growth in real GDP of about 6 per cent per year, and significantly more rapid growth in industry and service of 5.0 and 5.4 per cent per year compared to growth in the agricultural sector of 2.8 per cent. As a result, agriculture's share of GDP fell further to about 15 per cent. Agriculture's share of employment also fell as the non-agricultural sector absorbed all of the increases in the labour force between 2006 and 2016.

Table 1: Real GDP in Viet Nam, 2006–16

Year	Total	Agriculture	Industry	Services
2006	1,699,501	355,831	649,657	694,013
2007	1,820,667	369,905	697,499	753,263
2008	1,923,749	387,262	726,329	810,158
2009	2,027,591	394,658	769,733	863,200
2010	2,157,828	396,576	693,351	797,155
2011	2,292,483	413,368	746,069	856,691
2012	2,412,778	425,446	801,217	914,177
2013	2,543,596	436,642	841,953	975,592
2014	2,695,796	451,659	896,042	1,035,726
2015	2,875,856	462,536	982,411	1,101,236
2016	3,054,470	468,813	1,056,808	1,178,143
Annual growth (%)	6.04	2.80	4.99	5.43

Notes: GDP in constant prices (billion 2010 VND). Sector GDP numbers do not add to total after 2009. The difference is production taxes and subsidies, which are reported separately.

Source: authors' compilation based on data from the General Statistical Office.

Table 2: Employment by sector

Year	Number of workers				Share of employment (%)		
	Agriculture	Industry	Services	Total	Agriculture	Industry	Services
2005	23,563.2	7,524.0	11,687.7	42,774.9	55.1	17.6	27.3
2006	23,747.4	8,044.6	12,199.5	43,991.4	54.0	18.3	27.7
2007	23,931.5	8,565.2	12,711.2	45,207.9	52.9	18.9	28.1
2008	24,303.4	8,985.5	13,171.9	46,460.8	52.3	19.3	28.4
2009	24,606.0	9,561.6	13,576.1	47,743.7	51.5	20.0	28.4
2010	24,279.0	10,277.0	14,492.5	49,048.5	49.5	21.0	29.5
2011	24,362.9	10,718.8	15,270.2	50,351.9	48.4	21.3	30.3
2012	24,357.2	10,896.5	16,168.7	51,422.4	47.4	21.2	31.4
2013	24,399.3	11,086.0	16,722.5	52,207.8	46.7	21.2	32.0
2014	24,408.7	11,229.0	17,106.8	52,744.5	46.3	21.3	32.4
2015	23,259.1	12,018.0	17,562.9	52,840.0	44.0	22.7	33.2
2016	22,315.2	13,199.0	17,788.6	53,302.8	41.9	24.8	33.4

Notes: estimates for 2006 are interpolated using 2005 and 2007. Estimates for 2016 are preliminary.

Source: authors' compilation based on data from the General Statistical Office.

3 Data

We use biennial household panel data from the VARHS that covers most of Viet Nam and provides a panel of over 2,100 households for the period between 2006 and 2016. The VARHS includes households from 12 provinces drawn from both north and south Viet Nam.[3] Our analysis focuses on the cropping sector, which generated the bulk of farm income for rural households throughout the period. We supplement the VARHS with data from the Vietnam Household Living Standards Survey (VHLSS).

Our analysis requires detailed measures of inputs and outputs for all farming households during the period. To construct measures of productivity at the household farm level, we need information on crop output as well as farm inputs in the form of cultivated land, capital, labour, and intermediate inputs such as fertilizers. A household farm may operate multiple plots of land and produce more than one type of crop. Since the production unit for our analysis is the household farm, we aggregate all outputs and inputs at the household farm level. Below we describe the construction of these variables. Additional details on the data construction are provided in the Appendix.

3.1 Real gross output

We construct a measure of real gross output at the farm level by aggregating physical production of each crop using a set of common crop prices for all households. These prices are computed using a combination of household-reported information on sales quantities and revenues, and the estimated value of their harvest when none of the output is sold. The price for each crop is constructed as an average of the median annual price between 2006 and 2016. This procedure reduces differences in the value of gross output between households to those arising primarily from differences in the quantity of crops produced, and is crucial for productivity measurement.

[3] The regions include the Red River Delta, Northeast, Northwest, and North Central Coast in the North and the South Central Coast, Central Highlands, and Mekong Delta in the South. The survey does not contain households from the Southeast region.

3.2 Capital

The stock of capital is constructed as the sum of three types of farm capital: (1) non-durable capital; (2) durable capital; and (3) capital services. Non-durables are reported by the household and include small equipment (e.g. sickles). Durable capital consists of machinery and equipment owned by the household, such as pesticide sprayers, water pumps, and tractors. Using a set of common prices to value this equipment, and information on the age of this machinery to estimate depreciation, we construct estimates of the value of durable capital. Finally, each household reports expenditures on hired-in capital services (e.g. rental of equipment). To convert expenditure on capital services into a stock we use a risk-free interest rate (assumed to be equal to the return on one-year Viet Namese government bonds) and a depreciation rate of 8 per cent. The stock is then adjusted for differences in the price of capital services between years using a price deflator we construct from information on the household supply of capital services in the VHLSS.

3.3 Land

Plots of land and their use are reported by the household. We measure land as cultivated area operated by the household for the growing of annual and perennial crops. Plots used for other activities (e.g. forestry) are not included. Our land measure includes land that is owned and operated by the household as well as land that is rented in. It excludes land that is rented out.

3.4 Labour

Labour input on the farm is constructed from two sources: (1) labour supplied by members of the household; and (2) labour hired by the household. Household labour is the sum of the number of days worked by household members in activities related to the growing of crops. Households do not report labour hired in terms of days, but rather total expenditure. To convert expenditure into days, we construct a daily wage rate implied by household income and time worked in agriculture outside of the family farm.

3.5 Intermediate inputs

Expenditure on intermediate inputs is reported by the household for a list of goods including seeds, fertilizers, and pesticides, of which expenditure on fertilizer consistently represents over half. Lacking price data for other intermediates, we deflate expenditure on fertilizers using a region–year fertilizer price index constructed from the VHLSS data to obtain an estimate of real intermediate input use.

3.6 Sample selection

We restrict our sample to 2,118 households that we observe in the survey in all six years. While our focus is the agricultural sector, this sample contains households that enter, exit, and never participate in crop production. In addition, we drop a small number of outliers from the entire balanced panel on the basis of output-per-land and output-per-worker.[4] After this additional trimming, the final sample contains a balanced panel of 2,087 households.

4 Big picture

Drawing on the balanced panel of households in the VARHS data (with 2,087 household observations), Table 3 captures aggregate changes between 2006 and 2016 in inputs and outputs in the cropping sector in agriculture. These data suggest fairly rapid growth in the gross value of crop output, which grew at a real annual rate of 4.4 per cent and in total by 53 per cent over the 10-year period.[5] Aggregate numbers conceal important differences by crop, and changes in the structure of production. Growth in rice and other cereals (primarily maize) lagged increases in annual crop and perennials such as coffee, cocoa, cashews, and pepper. As a result, cereals' share of real output declined by 15 percentage points from 59.9 to 44.2, while perennials' share rose from 28.7 to 44.2 per cent (see Table 4).

Table 3: Farm output and input growth, 2006–16

Year	Farm output	Capital	Labour	Land	Intermediates	TFP
2006	100.0	100.0	100.0	100.0	100.0	100.0
2008	113.0	79.7	109.5	104.7	104.2	108.9
2010	124.9	62.6	91.7	101.5	103.2	129.6
2012	125.6	65.2	78.4	101.7	110.5	133.5
2014	151.3	65.2	65.3	96.2	109.0	172.6
2016	153.4	70.7	65.3	91.0	121.1	170.8
Annual growth (%)	4.4	–3.4	–4.2	–0.9	1.9	5.5

Notes: values in 2006 normalized to 100. All values based on common price indices. See Section 3 for details on construction of values.

Source: authors' own calculations based on data from VARHS.

Table 4: Real output growth by crop, 2006–16

	Share 2006 (%)			Share 2016 (%)			Growth 2006–16 (%)		
	Viet Nam	North	South	Viet Nam	North	South	Viet Nam	North	South
Cereals	59.9	77.1	51.9	44.2	72.2	37.5	13	–14	32
Rice	52.0	62.2	47.3	39.8	57.3	35.7	18	–15	38
Maize	6.4	12.0	3.8	3.3	11.6	1.3	–21	–10	–37
Other	1.5	2.9	0.8	1.1	3.2	0.5	9	4	18
Annuals	4.7	5.3	4.5	6.4	15.5	4.2	107	169	72
Vegetables	3.7	5.5	2.9	1.0	2.8	0.5	–61	–54	–67
Fruits	2.9	4.5	2.2	4.3	3.7	4.4	126	–23	273
Perennials	28.7	7.6	38.7	44.2	5.9	53.3	136	–29	151

Notes: change in real output of crops from 2006 to 2016. Other cereals includes potatoes, sweet potatoes, and cassava. The value of crop output is computed using a set of common prices.

Source: authors' own calculations based on data from VARHS.

Table 3 also provides information aggregated up from the balanced household panel on input use, which includes labour, capital, land, and intermediates. Especially noteworthy in Table 3 is the sharp decline

[4] Specifically, we drop households for whom output-per-land is ±4 log points from the mean and output-per-land is ±5 log points from the mean. We drop these households from all six years to maintain the balanced panel.

[5] Estimates for the growth in real output in the balanced panel are higher than the 3 per cent annual growth suggested by the national data (Table 1). Differences in regional coverage, as well as possible problems in national price deflators, may be responsible for these differences.

in labour input. This occurs on both the extensive and intensive margins: some households are exiting agriculture, while among those who continue to farm, labour input is also declining. Table 5 breaks down labour supply to agriculture in more detail. Over this period, total labour supply to agriculture by households in our panel declines by almost 40 per cent from 161 days per year to 96 days. Contributing nearly equally to this reduction is a decline in the number of individuals working in agriculture and a decline in the number of days worked by those who continue to work in agriculture.[6] Underlying the decline in labour supply to agriculture is expanded off-farm opportunities, especially in the secondary (manufacturing and construction) sector.[7]

Table 5: Household labour supply

Year	2006	2008	2010	2012	2014	2016	Annual change (%)
Number of households	2087	2087	2087	2087	2087	2087	–
Number farming	1823	1828	1782	1614	1699	1582	–1.4
Percentage in farming	87.4	87.6	85.4	77.3	81.4	75.8	–
Average household size	4.60	4.59	4.38	4.26	4.17	4.06	–1.3
Average number working	3.07	3.10	3.01	2.83	2.88	2.78	–1.0
Working in agriculture	2.60	2.48	2.47	2.23	2.22	2.04	–2.5
Working in non-agriculture	1.70	2.04	1.98	1.93	2.02	1.99	1.6
Self-employed	0.86	1.22	1.20	1.14	1.15	0.97	1.2
Wage worker	0.97	1.00	1.02	1.06	1.18	1.27	2.7
Average household labour supply							
Agriculture, own farm	161	172	141	116	99	96	–5
Non-agriculture	–	220	206	209	221	219	0
Self-employed	–	61	49	48	44	38	–5.7
Wage worker	–	160	157	161	178	181	1.6

Notes: household labour supply calculated on balanced panel data. Number of farms is calculated as the number of households that report positive production of crops in a given year. Working in agriculture reports the number of household members actively working on the household's farm. Working in non-agriculture includes all household members actively working outside the household's farm. This includes members working outside of the household in the agricultural sector. Self-employed includes all members not working for a wage and not working on the household's farm. Wage workers includes all workers employed at a wage outside of the household. Total household labour supply is measured in units of effective days. Information on labour supplied for non-agriculture, self-employment, and wage work is unavailable for 2006. Annual change calculated as $(X_{2016}/X_{2006})^{1/10}$.

Source: authors' own calculations based on data from VARHS.

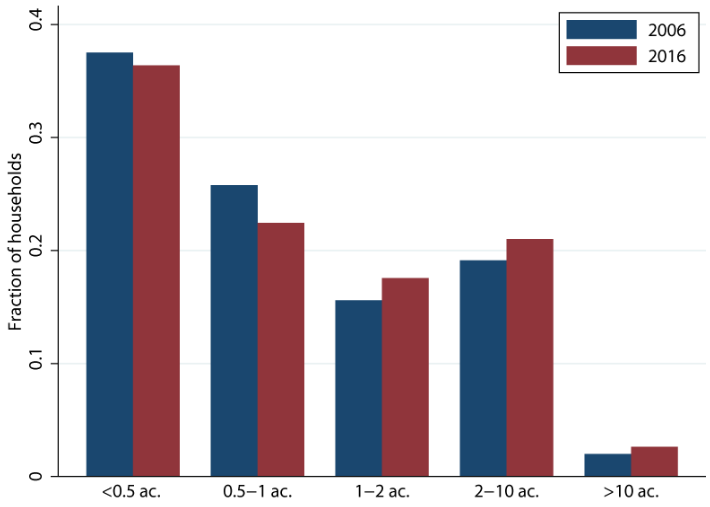
Table 6 suggests much more modest changes with respect to land. Cultivated and sown land both decline slightly. MCI, our measure of cropping intensity, is nearly identical at the beginning and end of the period. Several changes are noteworthy, however. First, the decline in cultivated area is accompanied by a slight shift to the right in the distribution of farm size. Figure 1 shows the changes in the distribution of farm size measured in terms of cultivated land. Farms of less than one acre decline slightly, while those larger than one acre increase. Second, in the context of a decline in the total number of plots farmed, average plot size increased from 1,396 m^2 to 1,850 m^2, an increase of one-third. Finally, households in our panel, on net, go from renting in land to renting out. In 2006, households rented-in 8.1 per cent of the land they farmed compared to the 3.5 per cent they rented out; in 2016, they rented-in 7 per cent, but

[6] Days in the cropping sector represent about two-thirds of total labour supply to farming, with the rest in animal husbandry, aquaculture, and forestry. Over this period, number of days in these non-cropping activities decline commensurately with that in the cropping sector.

[7] The survey may actually underestimate the shift to non-agriculture. Individuals who migrate and reside outside the home for more than six months of the year are not classified as household members. Reallocation of labour within the household associated with increases in longer-term migration would not be captured.

rented out 11 per cent. This shift is likely tied to the expanding role of non-agricultural activity among these households.

Figure 1: Land-size distribution of farms, 2006 and 2016



Notes: land-size calculated as cultivated land used in the production of crops.

Source: authors' calculations based on data from VARHS.

Table 6: Household land

		2006	2008	2010	2012	2014	2016	Annual change (%)
Totals								
Owned land	(000 m^2)	12,852	14,019	13,622	13,794	13,102	12,855	0
Cult. area	(000 m^2)	13,463	14,289	13,660	13,836	12,879	12,365	–0.9
Sown area	(000 m^2)	21,095	20,357	20,518	20,377	18,591	19,254	–0.9
MCI	(000 m^2)	1.57	1.42	1.50	1.47	1.44	1.56	0
Number of plots	(000 m^2)	9,980	9,867	9,408	9,007	7,880	7,416	–2.8
Rental rates								
Rent in (%)	(000 m^2)	8.1	6.7	6.8	6.6	5.8	7.0	–
Rent out (%)	(000 m^2)	3.5	4.8	6.3	6.3	7.5	11.0	–
Farming households:								
Avg. cult. area	(m^2)	6,457	6,846	6,544	6,633	6,170	5,925	–0.9
Avg. plot size	(m^2)	1,396	1,517	1,546	1,632	1,757	1,850	2.9%

Notes: cultivated land is the sum of land area in thousands of square metres of plots used for crop production by households in the balanced panel. It includes both land owned by the household as well as land rented in. Owned land includes both land owned and used by the household and land rented out. The multi-cropping index (MCI) is calculated as the average number of seasons cultivated per plot weighted by plot area. Due to missing data, sown area cannot be directly calculated over the entire sample. To correct for data limitations, sown area reported is calculated as cultivated land multiplied by the MCI.

Source: authors' own calculations based on data from VARHS.

Table 7 provides complementary information on sown area by crop in 2008 and 2016 for all Viet Nam and for the two regions separately. Over this period, total sown area declined by 5 per cent; however all of this occurred in the north, primarily in the cultivation of rice. In the south, sown area actually increased slightly as area in cash crops such as fruits and perennials offset the reduction in cereals and other annual crops.

Table 7: Growth in sown area by crop, 2008–16

	Share 2008 (%)			Share 2016 (%)			Growth 2008–2016 (%)		
	Viet Nam	North	South	Viet Nam	North	South	Viet Nam	North	South
Cereals	72.7	86.4	64.1	69.0	86.0	60.4	-9.6	–17.1	–3.2
Rice	59.5	66.9	54.9	60.0	69.0	55.5	–4.0	–14.3	3.8
Maize	9.2	14.3	6.0	7.2	13.4	4.0	–25.9	–22.2	–31.4
Other	4.0	5.1	3.2	1.9	3.7	0.9	–55.5	–40.3	–70.5
Annuals	2.3	1.8	2.6	1.8	1.8	1.8	–25.6	–14.8	–30.3
Vegetables	1.8	2.6	1.4	1.9	3.4	1.2	1.1	10.0	–9.3
Fruits	2.9	3.1	2.8	4.7	2.1	6.0	53.3	–43.4	120.9
Perennials	20.3	6.1	29.1	22.5	6.6	30.6	5.6	–10.0	7.7
Total	100.0	100.0	100.0	100.0	100.0	100.0	–4.8	–16.8	2.6

Notes: estimates are based on data for the balanced panel. Sown area is calculated as the area of land cultivated multiplied by the number of seasons the land is actively used. 2006 is not included because of data issues.

Source: author's own calculations based on data from VARHS.

4.1 Growth accounting

Drawing on aggregate data for the balanced panel, we examine the implications of changes in inputs and outputs for productivity in agriculture. In our growth accounting exercise, we assume an aggregate production function for agriculture of the form:

$$Y_{a,t} = A_{a,t}\left[\left(K_{a,t}^{\alpha}L_t^{\beta}N_{a,t}^{1-\alpha-\beta}\right)^{1-\theta}M_t^{\theta}\right]^{\gamma}I_t^{1-\gamma}, \tag{1}$$

where $Y_{a,t}$ is aggregate real crop output at date t, $K_{a,t}$ is aggregate capital in agriculture at date t, L_t is aggregate land input in agriculture, $N_{a,t}$ is aggregate labour days in agriculture at date t, M_t is aggregate real intermediate inputs used in agriculture at date t, and I_t is the number of farms in agriculture at date t. Total factor productivity (TFP) $A_{a,t}$ is calculated as a residual by subtracting from output the contribution of measured inputs given by the production function in equation (1). Note that this aggregate production function can be obtained from the aggregation of household-level farm production functions we consider in our analysis in Section 5.[8]

Using the production function in equation (1) and values for $\alpha = 0.09$, $\beta = 0.36$, $\theta = 0.35$, and $\gamma = 0.85$, we decompose the growth in real gross output over time into that which can be attributed to growth in factor inputs and TFP.[9] We report the results of this growth accounting in Table 8. Overall, the growth accounting exercise suggests that all of the growth in real gross crop output is coming from improvements in productivity. Increases in the use of intermediate inputs are more than offset by reductions in labour input use, as well as land and capital. Productivity growth over this period averages more than 5 per cent per annum.

There are also important regional differences in the growth process in agriculture. In order to highlight these differences, we report the growth accounting for the period between 2006 and 2016 separately for the north and south in Table 9. Significant regional differences in the growth of output emerge. In fact, all of the growth in crop output is occurring in the south; in the north, the real value of crop output actually declines slightly. These differences in growth are a product of differences in input use and

[8] See Adamopoulos et al. (2017) for a detailed characterization of the aggregate production function from farm-level production.

[9] We discuss in detail the selection of parameter values and their implications for input shares in Section 6.

productivity growth. In the case of TFP, growth in the south is almost two times higher than it is in the north. This is reinforced by much smaller reductions in input use in the south.[10]

Table 8: Growth accounting, 2006–16

Year	Output (%)	Capital (%)	Labour (%)	Land (%)	Intermediates (%)	TFP (%)
Growth						
2006–08	13.1	–20.3	9.5	4.7	4.2	8.9
2008–10	10.5	–21.5	–16.3	–3.0	–1.0	19.0
2010–12	0.6	4.1	–14.5	0.2	7.1	3.0
2012–14	20.6	0.0	–16.7	–5.4	–1.4	29.3
2014–16	1.4	8.5	0.0	–5.4	11.1	–1.1
2006–16	4.4	–3.4	–4.2	–0.9	1.9	5.5
Contribution						
2006–16	–	-4.0	–29.9	–4.4	13.4	124.9

Notes: the growth accounting is based on the aggregate production function for agriculture given by equation (1). Contribution refers to the percentage of output growth that is accounted for by each factor input and TFP. Outputs and inputs are aggregated from household farm-level information. TFP ($=A_{a,t}I_t^{1-\gamma}$) is calculated using equation (1).

Source: authors' own calculations based on data from VARHS.

Table 9: Growth accounting 2006–16: regional differences

	Output (%)	Capital (%)	Labour (%)	Land (%)	Intermediates (%)	Residual (%)
National	4.4	–3.4	–4.2	–9.0	1.9	5.5
North	–0.8	–2.7	–6.5	–2.6	–3.6	3.0
South	6.2	–3.9	–2.0	0.0	4.5	5.6

Notes: the table reports annualized growth in values. Based on aggregate production function in equation (1) with parameter values $\alpha = 0.0769$, $\beta = 0.3077$, $\theta = 0.2353$, and $\gamma = 0.85$. TFP ($=A_{a,t}I_t^{1-\gamma}$) calculated using equation (1).

Source: authors' own calculations based on data from VARHS.

5 Framework for measuring misallocation

A potential source of the aggregate productivity gains identified in Table 8 is improvements in resource allocation in the farm sector. Changes in resource allocation may also underlie the differences between the two regions. In this section, we describe our framework for measuring misallocation in Viet Nam agriculture. We characterize the efficient allocation—the allocation across a fixed set of farmers that maximizes agricultural output given total factor inputs—and compute two measures of misallocation: the standard deviation of log total factor productivity revenue (TFPR), and the total factor productivity gains of reallocating resources from the actual to the efficient allocation.

5.1 Description

Consider an agricultural economy in which a single output is produced by a set of production units. The production unit is a farm that is operated by heterogeneous farmers, indexed by i, with farming ability s_{it} in period t. There is a fixed number I_t of farmers in period t. A farmer with ability s_{it} produces according

[10] These regional comparisons conceal significant differences within the north and south that we do not pursue here. In the north we find significant productivity growth in the Red River Delta, and much lower, if not negative, TFP growth in the other regions.

to a decreasing returns to scale technology:

$$y_{it} = s_{it}^{1-\gamma} \left[\left(k_{it}^{\alpha} \ell_{it}^{\beta} n_{it}^{1-\alpha-\beta} \right)^{1-\theta} m_{it}^{\theta} \right]^{\gamma}, \tag{2}$$

where y_{it}, k_{it}, ℓ_{it}, n_{it}, and m_{it} denote real farm gross output, capital input, land input, labour input, and intermediate inputs such as fertilizer. The parameter $\gamma \in (0,1)$ is the span-of-control of the farmer, which describes the extent to which more productive farmers can manage more resources. We refer to s_{it} as farming ability and $s_{it}^{1-\gamma}$ as farm-level productivity or TFP.

Note that a key feature of this production function is that it is not optimal to allocate all inputs to the most productive farmer since there are decreasing returns to the allocation of inputs to any given farmer. This implies a non-degenerate distribution of farm sizes which we observe in reality. The parameters $\alpha, \beta, \theta \in (0,1)$ capture the relative importance of capital, land, labour, and intermediates in production. Also note that factor inputs correspond to the amount of inputs used in production rather than inputs owned by the farmer. For example, in our analysis what matters is the operational scale of the farm measured by the amount of cultivated land of the farm rather than the amount of land owned by the farmer.

Given this production structure, actual aggregate agricultural output in the economy is

$$Y_t = \sum_{i=1}^{I_t} y_{it}. \tag{3}$$

5.2 Efficient allocation

In the context of the preceding economic environment, we now define the efficient allocation as the allocation that maximizes aggregate agricultural output. Formally, the efficient allocation solves the planner's problem given by,

$$Y_{it}^{e} = \max_{\{k_{it}, \ell_{it}, n_{it}, m_{it}\}_{i=1}^{I_t}} \sum_{i=1}^{I_t} s_{it}^{1-\gamma} \left[\left(k_{it}^{\alpha} \ell_{it}^{\beta} n_{it}^{1-\alpha-\beta} \right)^{1-\theta} m_{it}^{\theta} \right]^{\gamma}, \tag{4}$$

subject to the resource constraints

$$K_t = \sum_{i=1}^{I_t} k_{it}; \quad L_t = \sum_{i=1}^{I_t} \ell_{it}; \quad N_t = \sum_{i=1}^{I_t} n_{it}; \quad M_t = \sum_{i=1}^{I_t} m_{it},$$

where K_t, L_t, N_t, and M_t are the aggregate amounts of capital, land, labour, and intermediates used in production in period t. This problem assumes that aggregate factor inputs are fixed and that the allocation is over a fixed set of existing farmers.

The efficient allocation is easy to characterize from the above problem. Factor inputs are allocated to where productivities are highest, resulting in equalization of marginal products across all producers, and the maximization of output. In the efficient allocation, resources are strictly linked with relative farming ability s_{it}, with more productive farmers allocated more of each input. In particular, for any factor input $x \in \{k, l, n, m\}$:

$$x_{it}^{e} = \frac{s_{it}}{\sum_{j=1}^{I_t} s_{jt}} X_t, \tag{5}$$

where x_{it}^{e} is the efficient allocation of factor X to household i at time t.

In the efficient allocation (5), aggregate agricultural output is given by,

$$Y_t^e = (\bar{S}_t I_t)^{1-\gamma} \left[\left(K_t^{\alpha} L_t^{\beta} N_t^{1-\alpha-\beta} \right)^{1-\theta} M_t^{\theta} \right]^{\gamma}, \tag{6}$$

where $\bar{S}_t = \frac{1}{I_t} \sum_{i=1}^{I_t} s_{it}$.

5.3 Identification of distortions

The observed allocation of resources differs from the efficient allocation that solves the problem described by equation (4). In the efficient allocation, the marginal products of factors are equalized across farms; in the observed distorted allocation, this is not the case and marginal products differ across farms. With misallocation, there is a potential gain from reallocating resources from low marginal product to high marginal product farmers. For example, a high-productivity farmer that is unable to acquire additional land to operate would have a high marginal product of land. Regional and time variation in these types of institutions are discussed in Section 7.

Distortions can be measured at the individual factor-market level; for example, for markets in land, intermediate inputs, capital, and labour. In this paper, we focus on a composite measure of distortions at the farm level, TFPR, that aggregates distortions over all four production factors:

$$\begin{aligned} \text{TFPR}_{it} &= \frac{y_{it}}{\left(k_{it}^{\alpha} \ell_{it}^{\beta} n_{it}^{1-\alpha-\beta} \right)^{1-\theta} m_{it}^{\theta}} \\ &= \left(\left(\frac{\text{MPK}_{it}}{\alpha(1-\theta)\gamma} \right)^{\alpha} \left(\frac{\text{MPL}_{it}}{\beta(1-\theta)\gamma} \right)^{\beta} \left(\frac{\text{MPN}_{it}}{(1-\alpha-\beta)(1-\theta)\gamma} \right)^{1-\alpha-\beta} \right)^{1-\theta} \left(\frac{\text{MPM}_{it}}{\theta\gamma} \right)^{\theta}, \end{aligned} \tag{7}$$

where MPX_{it} is the marginal product of factor X for farm i in period t. TFPR_{it} in equation (7) is a composite measure based on the marginal products at the firm level and in the efficient allocation is the same for all farms. In this regard, dispersion in TFPR_{it} is a measure of the allocative efficiency of the economy with greater dispersion indicating larger distortions and inefficiency. In the efficient allocation, there is no variation in TFPR_{it}.

6 Productivity and misallocation

We use our simple framework and the data for Viet Nam to measure farm productivity and characterize misallocation in agriculture across farms, across regions, and over time.

6.1 Measuring farm productivity

The first step in characterizing misallocation in Viet Nam agriculture is to estimate TFP at the farm level. On the basis of the production function in equation (2), we measure productivity as the ratio of output to inputs, or

$$s_{it}^{1-\gamma} = \frac{y_{it}}{\left[\left(k_{it}^{\alpha} \ell_{it}^{\beta} n_{it}^{1-\alpha-\beta} \right)^{1-\theta} m_{it}^{\theta} \right]^{\gamma}}, \tag{8}$$

where $s_{it}^{1-\gamma}$ is farm TFP. We construct measures of gross farm output and inputs using the data described in Section 3.

To calibrate the parameters of the production function, we use a span-of-control $\gamma = 0.85$; capital share $\alpha(1-\theta)\gamma = 0.05$; land share $\beta(1-\theta)\gamma = 0.20$; unskilled labour share $(1-\alpha-\beta)(1-\theta)\gamma = 0.30$; and intermediates share $\theta\gamma = 0.30$. Factor shares are the same as those used earlier in the growth accounting exercise and are based on the income shares for capital, labour, and intermediates in farming in our household data.[11] These data suggest a capital share of 5 per cent, a labour share of 45 per cent, and an intermediate share of 35 per cent. We target an intermediate share of 30 per cent as a compromise with estimates from other studies. For labour, we allocate one-third, or 15 per cent of total factor returns, to the management of the farm and 30 per cent to unskilled labour (supplied by household and non-household members). The share of land is computed as a residual and set equal to 20 per cent.

Aggregation of activity on multiple plots to the farm level helps attenuate concerns about unmeasured shocks and measurement error. Potential measurement error remains, however. To further mitigate these concerns, we divide our panel into two sub-periods of three rounds each, 2006–2008–2010 and 2012–2014–2016, and average household inputs and outputs in each of those two sub-periods. For example, for a household farm that operates in all three rounds within a sub-period, capital is calculated as the average capital in the three rounds. We do similarly for output and all other inputs. We also trim the top and bottom 1 per cent of observations ranked by farm TFP in each year to remove the influence of outliers.

We document the resulting distributions of farm productivity (TFP) in Figure 2 and in Tables 10. As a summary statistic of the dispersion in farm-level TFP, the standard deviation of log TFP is 0.46 in 2006 and 0.59 in 2016, reflecting both substantial dispersion in farm TFP as well as an increase over time. Using sub-period averages for the calculations, the dispersions in farm TFP in 2006–10 and 2012–16 are similar to our estimates for 2006 and 2016, as is the increase over time. Table 10 also reports the TFP ratio at different ranges of the productivity distribution. The TFP ratio between farms in the 99th and 1st percentile is 2.22 in 2006–10 and widens to 2.74 in 2012–16; for the 95th to 5th percentiles, the ratio rises from 1.57 in 2006–10 to 1.88 in 2012–16. These estimates of the dispersion in farm-level TFP are broadly consistent with estimates of micro-level productivity in other settings.[12] Table 10 also reports statistics separately for the two regions. Figure 3, on the other hand, captures the evolution of the productivity distribution in the north and south. In both regions an increase in productivity—reflected by a rightward shift in the distribution—is accompanied by widening in dispersion. Moreover, in the north we observe both a smaller increase in productivity and a larger increase in dispersion.

Table 10: Distribution of farm TFP

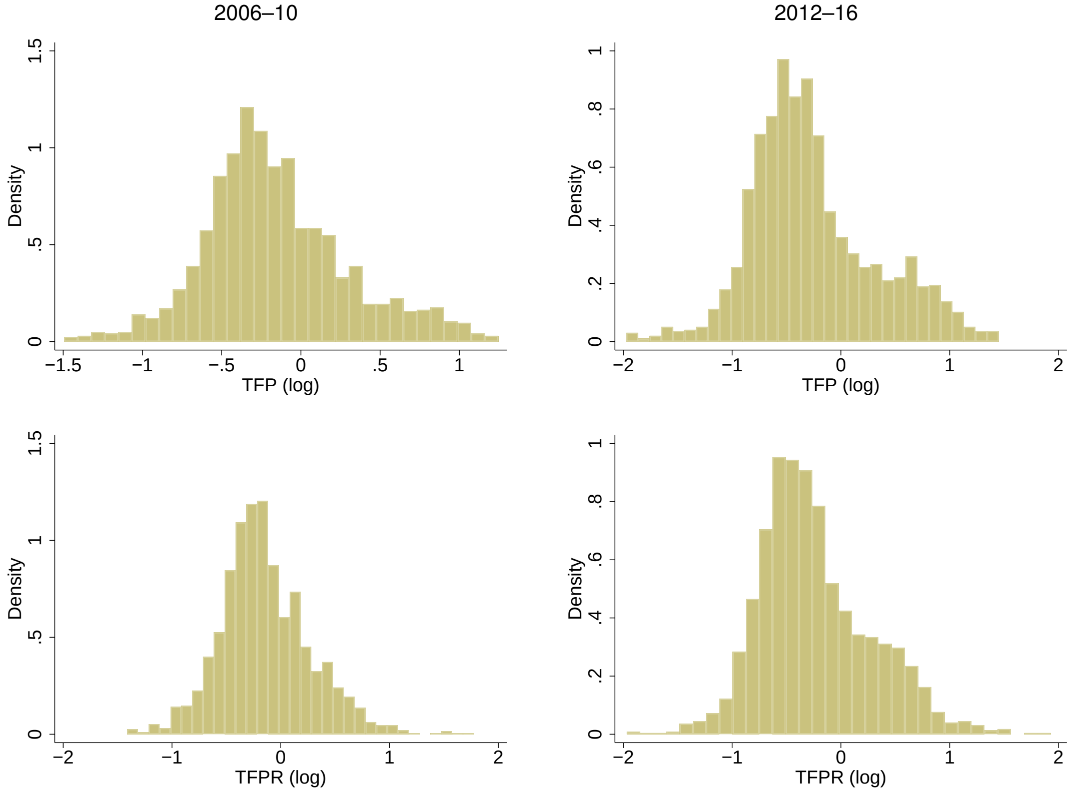
	National			North			South		
	Std dev.	99–1	95–5	Std dev.	99–1	95–5	Std dev.	99–1	95–5
2006–10	0.46	2.22	1.57	0.32	1.79	1.06	0.48	2.33	1.67
2012–16	0.58	2.74	1.88	0.40	2.42	1.22	0.57	2.82	1.80

Notes: distribution of TFP by period and region. Std dev. reports the standard deviation of $\log(\text{TFP}_{it})$. 99–1 reports the difference in $\log(\text{TFP}_{it})$ between the 99th percentile and 1st percentile household. 99–5 reports the difference in $\log(\text{TFP}_{it})$ between the 95th percentile and 5th percentile household. All numbers based on the final balanced panel of households with the top and bottom 1 per cent of observations trimmed based on TFP. Farm-level TFP corresponds to $s_{it}^{1-\gamma}$ in the model.
Source: authors' own calculations based on data from VARHS.

[11] We calculate the income shares as aggregate nominal expenditure on each factor divided by nominal output. We take nominal expenditure on intermediates to be equal to the expenditure on intermediates reported by households. We calculate a nominal wage rate using a Mincer regression on individual characteristics and then value labour supplied within households for the production of crops. We calculate the cost of capital using an interest rate equal to the sum of the one-year Viet Namese government bond rate and a depreciation rate of 8 per cent.

[12] For example, Restuccia and Santaeulalia-Llopis (2017) for farms in Malawi; Adamopoulos et al. (2017) for farms in China; Hsieh and Klenow (2009) for manufacturing plants in China, India, and the United States.

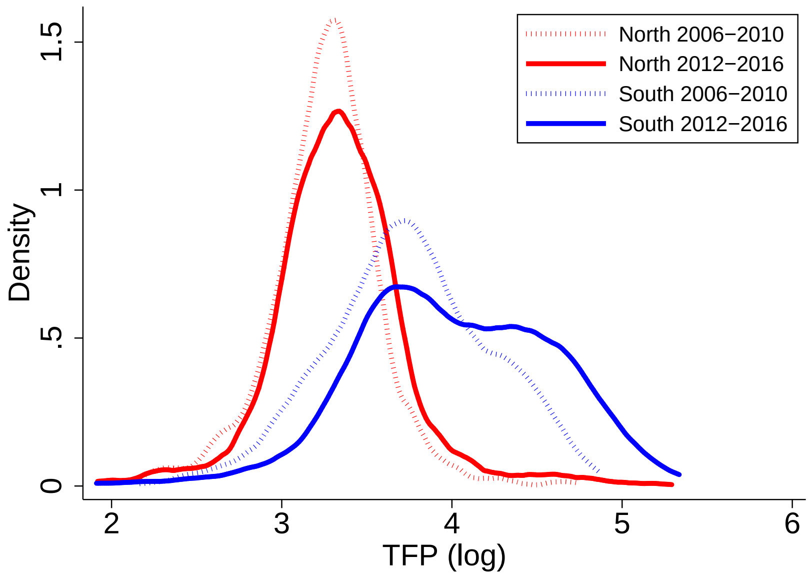
Figure 2: Distribution of farm-level TFP and TFPR



Notes: histograms for farm distribution of TFP and TFPR in periods 2006–10 and 2012–16. Each observation is a farm-period value of TFP or TFPR. Average TFP and TFPR are normalized to 1 in each period. Farm-level TFP corresponds to $s_{it}^{1-\gamma}$ in the model.

Source: authors' calculations based on data from VARHS.

Figure 3: Farm-Level TFP across time and region



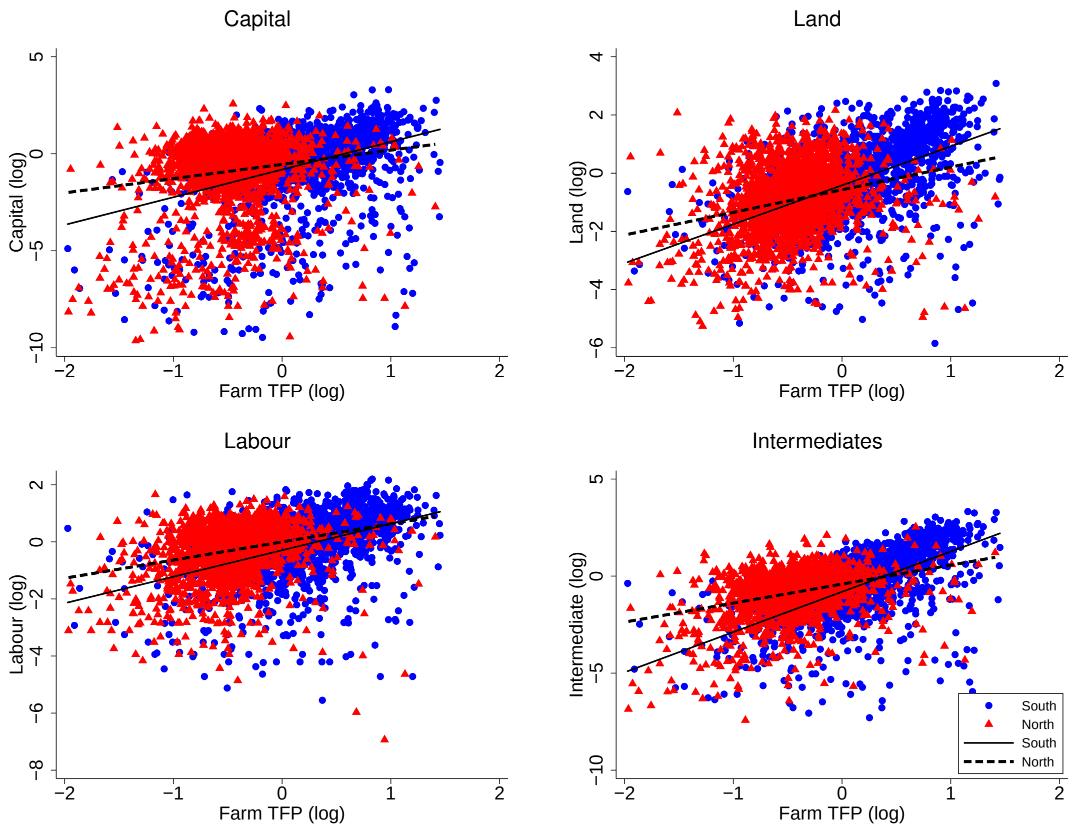
Notes: kernel density estimate of the distribution of TFP across farms in each period. Farm-level TFP corresponds to $s_{it}^{1-\gamma}$ in the model.

Source: authors' calculations based on data from VARHS.

6.2 Factor allocations, productivity, and misallocation

Recall that in the benchmark efficient allocation of our basic framework, factor inputs are strictly related to farm TFP. In Figure 4 we report input use by each farm in relation to their TFP for the two periods. In the figure, each dot represents a farm observation and the line is the fitted average relationship between farm input and TFP for both the north and south. A prominent characteristic of these allocations is that the elasticity of factor inputs with farm TFP is weak, but especially so in the north. For instance, in the north the elasticities for land, labour, and capital are between 0.64 and 0.98, but in the south are significantly higher and between 0.93 and 2.07. In the efficient allocation in which higher-productivity farms are allocated more factor inputs, the elasticity (slope) would be much higher (approximately 6.7).[13] There is also substantial dispersion in factor inputs among farms with the same TFP. These patterns—low correlation of factor inputs with farm TFP and dispersion in inputs within TFP types—reflect misallocation of factor inputs across farms, with the pattern of misallocation stronger in the north than in the south.

Figure 4: Factor inputs and farm-level TFP



Notes: each point in the figure is a farm-period observation. TFP and factor inputs are normalized to 1 in each period.
Source: authors' calculations based on data from VARHS.

As a summary measure of misallocation, we report the dispersion of log TFPR across farms. Recall from our basic framework that in an efficient allocation, marginal products of each factor are equalized across farms, as would be the marginal product of the composite input. Thus, dispersion in the gross output per unit of composite input reflects distortions in our setting. Table 11 reports the dispersion in TFPR across farms in the two sub-periods as well as other moments of the distribution of TFPR. The dispersion of

[13] Note that $\text{TFP}_{it} = s_{it}^{1-\gamma}$. In the efficient allocation, farms receive factors proportional to productivity s_{it} (e.g. $k^e_{it} \propto s_{it}$) or equivalently proportional to $\text{TFP}_{it}^{\frac{1}{1-\gamma}}$. The elasticity (slope) in the efficient allocation for Figure 4 is $\frac{1}{1-\gamma} \approx 6.7$.

log TFPR is 0.42 in 2006–10 and 0.52 in 2012–16, and indicative of high and rising misallocation in Viet Nam agriculture.[14] The misallocation measure is also larger in the south than in the north by 10 percentage points in both sub-periods.

Deviations of marginal (or average in our setting) products across farms is a symptom of misallocation, but the productivity cost of misallocation depends not only on the distribution of TFP across farms, but more generally on the joint distribution of TFP and TFPR. We next compute the productivity cost of misallocation.

Table 11: Distribution of farm distortions

	National			North			South		
	Std dev.	99–1	95–5	Std dev.	99–1	95–5	Std dev.	99–1	95–5
2006–10	0.42	2.03	1.36	0.31	1.63	1.06	0.41	1.99	1.38
2012–16	0.52	2.46	1.65	0.37	2.17	1.10	0.48	2.53	1.43

Notes: distribution of TFPR by period and region. Std dev. reports the standard deviation of $\log(\text{TFPR}_{it})$. 99–1 reports the difference in $\log(\text{TFPR}_{it})$ between the 99th percentile and 1st percentile household. 99–5 reports the difference in $\log(\text{TFPR}_{it})$ between the 95th percentile and 5th percentile household. All numbers based on the final balanced panel of households with the top and bottom 1 per cent of observations trimmed based on TFP.

Source: authors' own calculations based on data from VARHS.

6.3 Gains from reallocation

We measure the productivity cost of misallocation as the counterfactual aggregate productivity gain from reallocating resources across farms from the actual to the efficient allocation in each sub-period. In principle, some of the productivity growth we documented could be a product of improvements in resource allocation that would be reflected in declining costs from misallocation. This does not appear to be the case. Our estimates of the productivity cost of misallocation in Table 12 for the full sample and for the north and south suggest relatively high and rising levels of misallocation. At the national level, an efficient reallocation of factor inputs across farms would produce an increase in aggregate agricultural TFP of 68 per cent in 2006–10 and 80 per cent in 2012–16. The reallocation gains are also much larger in the north than in the south: 86 per cent in the north and 43 per cent in the south in 2006–10 and 162 per cent in the north and 47 per cent in the south in 2012–16.

A significant portion of the inefficiency of resource allocation at the national level is coming from misallocation within regions as opposed to misallocation between regions.[15] More than 70 per cent of the national reallocation gains are realized through reallocating resources across farms within a region. As documented earlier, resource use is more positively correlated with productivity in the south than in the north and, indeed, reallocation gains are larger in the north than in the south. Nevertheless, in both regions the productivity cost of misallocation appears to be relatively high. Moreover, there does not appear to be any reduction in the degree and cost of misallocation. At the end of the period, the potential gains to improvement in resource allocation are at least as high as they were at the beginning of the period. In the north, reallocation gains almost double between 2006–10 and 2012–16.

[14] Restuccia and Santaeulalia-Llopis (2017) and Adamopoulos et al. (2017) find larger dispersion in farm distortions in Malawi and China. In each of these settings, for instance, operated land is essentially unrelated with farm TFP, whereas in Viet Nam this correlation is weak but not zero.

[15] For the analysis of within-region reallocation, we divide the sample into seven regions comprising the Red River Delta, Northeast, Northwest, North Central Coast, South Central Coast, Central Highlands, and the Mekong Delta.

Table 12: Aggregate productivity gains from reallocation (%)

	National		North		South	
Period	Total	Within	Total	Within	Total	Within
2006–10	68	45	86	59	43	38
2012–16	80	62	162	123	47	44

Notes: the table reports the aggregate productivity gains from an efficient reallocation of resources among the set of existing farmers and given the aggregate amount of factor inputs in each period. Gains are calculated as Y_t^e/Y_t where Y_t^e is output under the efficient allocation and Y_t is the observed actual output in period t. Since aggregate factor inputs are held fixed, output gains are equivalent to productivity gains.

Source: authors' own calculations based on data from VARHS.

Counterfactuals

Our analysis identifies two trends seemingly at odds. First, there is a large and robust growth in aggregate productivity over this period. Second, the gains from eliminating misallocation are rising over time. This occurs in the context of a widening of farm-level productivities over time (Table 10). We now consider counterfactual experiments designed to disentangle the contribution of changes in the distribution of farm-level productivities and changes in the distribution of farm-level distortions to changes in aggregate output and productivity.

Note that our measure of the cost of misallocation is static as it calculates the increase in aggregate output if resources are allocated efficiently, taking as given the aggregate resource endowment and the set of existing farm productivities. But over time there may be a change in aggregate resources and farm productivities in addition to changes in factor allocations. Moreover, because of potential changes in aggregate factors, it is no longer the case that aggregate output gains are equivalent to aggregate productivity gains. With this in mind, we decompose growth in aggregate output into three channels—(1) gains from factor accumulation; (2) gains from within-farm TFP growth; and (3) gains from resource reallocation—and compare these gains with that of aggregate efficient output in order to understand the changes in the cost of misallocation over time. The main goal of this exercise is to examine how changes in the distribution of farm-level productivities (channel 2) and changes in farm-level distortions (channel 3) contributed to the gains in output and productivity.[16]

To make the numbers comparable to the misallocation exercise, we use the samples constructed for the 2006–10 and 2012–16 periods. In this regard, the results are not directly comparable to the statistics reported for growth in other parts of the paper that are based on the biennial data. Additionally, we restrict the sample to include only households that are actively involved in crop production in both the 2006–10 and 2012–16 periods; hence, the misallocation numbers will differ slightly from the numbers reported earlier.[17]

First, the aggregate output gains from resource accumulation are calculated as

$$Y_{12-16}^{cf_1} = \sum_{i\in\mathscr{I}} \tilde{y}_{i,12-16}^{cf_1} = \sum_{i\in\mathscr{I}} s_{i,06-10}^{1-\gamma}\left[\left(\tilde{k}_{i,12-16}^{\alpha}\tilde{\ell}_{i,12-16}^{\beta}\tilde{n}_{i,12-16}^{1-\alpha-\beta}\right)^{1-\theta}\tilde{m}_{i,12-16}^{\theta}\right]^{\gamma},$$

[16] This analysis is analogous to productivity growth decompositions as in Foster et al. (2008), with the main difference that our aggregate measures of TFP follow the specific framework described in Section 5 and that, as a result, there is no closed form decomposition for the different channels.

[17] This analysis is based on 1,770 households that are active in agriculture in both the 2006–10 and 2012–16 periods. As mentioned previously, a household is considered active if it reports producing crops in any of the three surveys during the period.

where $\tilde{x} = x_{i,06-10}\frac{X_{12-16}}{X_{06-10}}$ is the counterfactual value of input x and X_t is the aggregate stock of input x in period t. The counterfactual $Y^{cf_1}_{12-16}$ is the aggregate output that would have been produced if only aggregate stocks of resources had changed between the periods, holding the allocation of resources fixed. Second, the aggregate output gains from within-farm TFP change are calculated as

$$Y^{cf_2}_{12-16} = \sum_{i \in \mathscr{I}} \tilde{y}^{cf_2}_{i,12-16} = \sum_{i \in \mathscr{I}} s^{1-\gamma}_{i,12-16} \left[\left(k^{\alpha}_{i,06-10} \ell^{\beta}_{i,06-10} n^{1-\alpha-\beta}_{i,06-10} \right)^{1-\theta} m^{\theta}_{i,06-10} \right]^{\gamma}.$$

The counterfactual $Y^{cf_2}_{12-16}$ is then the aggregate output that would have been produced if only farm-level productivities had changed between the periods. Third, the aggregate output gains from resource reallocation are calculated as

$$Y^{cf_3}_{12-16} = \sum_{i \in \mathscr{I}} \tilde{y}^{cf_3}_{i,12-16} = \sum_{i \in \mathscr{I}} s^{1-\gamma}_{i,06-10} \left[\left(\tilde{k}^{\alpha}_{i,12-16} \tilde{\ell}^{\beta}_{i,12-16} \tilde{n}^{1-\alpha-\beta}_{i,12-16} \right)^{1-\theta} \tilde{m}^{\theta}_{i,12-16} \right]^{\gamma},$$

where $\tilde{x}_{i,12-16} = x_{i,12-16}\frac{X_{06-10}}{X_{12-16}}$. The counterfactual $Y^{cf_3}_{12-16}$ is then the output that would have been produced if productivity remained unchanged and resources were reallocated to reflect the 2012–16 allocations. Note that the change in aggregate output from these three counterfactuals may not add to the actual change in output because of the potential interaction between these channels. As a result, we compute the difference between actual and counterfactual output growth as a residual.

At the national level, the increase in aggregate agricultural output in this sample between the two sub-periods is 30 per cent (see Table 13).[18] Changes in overall aggregate inputs in agriculture contribute negatively to this increase, an 8 percentage point reduction, as aggregate resource use declined between periods. Changes in within-farm TFP are the source of the bulk of the increase, a 32 per cent increase. The change in factor allocation among households contributes positively to this gain but plays a relatively small role, only a 3 per cent increase. The residual represents a 4 per cent increase. We observe similar patterns in the north and south regions with two salient differences: first, there is a much larger decline in factor accumulation in the north (16 percentage points) compared to the south (1 percentage point); and second, in both the main source of output growth is the increase in within-farm TFP, which is much larger in the south (41 per cent) than in the north (13 per cent).

In order to relate these counterfactuals with our misallocation results over time, we also compute the efficient output in each period and the efficient output in 2006–10 with 2012–16 farm productivity and with 2012–16 aggregate inputs. The results of these additional counterfactuals are reported in Table 13 for the entire economy (national) in each sub-period as well as for the north and south regions. In each case, output is calculated relative to actual output in 2006–10.

At the national level, the output gain of an efficient reallocation is 67 per cent in 2006–10 and 82 per cent in 2012–16. Efficient output in 2012–16 represents an overall increase relative to 2006–10 actual output of 136 per cent (a 2.36-fold increase). Even though improvements in resource reallocation between the two periods contribute positively to output growth, this reallocation is not strong enough to keep up with changes in farm-level TFP. As a result, there is an increase in the cost of factor misallocation over time. Intuitively, the cost of misallocation depends on the joint distribution of farm-level productivities and factor allocations, as opposed to only the distribution of factor allocations. To help illustrate this point, note that efficient output in 2006–10 with 2012–16 farm TFP is 2.56-times that of 2006–10 actual output. That is, despite holding farm-level factor allocations fixed, the cost of misallocation increases from 67 per cent in the 2006–10 period to 156 per cent when 2012–16 farm-level productivities are used.

[18] This growth is less than the total observed between 2006 and 2016, but in line with the expected change over a six-year period defined by the mid-points of the two periods.

Table 13: Counterfactual results on aggregate agricultural output

	National	North	South
Actual output			
2006–10 (Y_{06-10})	1.00	1.00	1.00
2012–16 (Y_{12-16})	1.30	0.98	1.45
Counterfactual output			
Factor accumulation ($Y^{cf_1}_{12-16}$)	0.92	0.84	0.99
Change in farm TFP ($Y^{cf_2}_{12-16}$)	1.32	1.13	1.41
Change in allocation ($Y^{cf_3}_{12-16}$)	1.03	1.01	1.01
Residual	1.04	1.02	1.03
Counterfactual efficient output:			
2006–10 (Y^{e}_{06-10})	1.67	1.75	1.43
with 2012–16 factors	1.54	1.47	1.41
with 2012–16 farm TFP	2.56	3.06	2.15
2012–16 (Y^{e}_{12-16})	2.36	2.57	2.13

Notes: output is calculated relative to agricultural output in 2006–10. In the counterfactual actual output, the residual is calculated as $\frac{Y_{16-16}}{Y^{cf_1}Y^{cf_2}Y^{cf_3}}$. Results are based on the balanced panel of households that are actively producing crops in both periods.

Source: authors' own calculations based on data from VARHS.

We highlight two conclusions from the counterfactual exercises. First, the majority of output growth over this period is driven by improvements in farm-level productivities, with a minor positive role for the reallocation of resources and a negative role for factor accumulation. However, this growth is uneven and is driven by an increase in the number of relatively high-productivity farms (Figure 3). This results in an increase in both the aggregate TFP and the dispersion of farm-level productivities over this period (Table 10). Second, while the allocation of resources improves over this period (see Y^{cf_2}), resources are not being reallocated as quickly as farm-level productivities are changing. The larger cost of misallocation in the period 2012–16 can be attributed to the changes in the joint distribution of farm-level productivities and factor allocations. Together, they explain the simultaneous rise in productivity and costs of misallocation.

7 Misallocation and institutional constraints

Our results suggest high and rising misallocation. The misallocation problem is more severe in the north compared to the south, and appears to have worsened over time in the north. Lower productivity growth in the north between 2006–16 may be associated with the same set of forces contributing to the increase in misallocation.

At its simplest, misallocation reflects the fact that resources are not being efficiently allocated across farms. And the likely sources are constraints on household choice and market imperfections. Making a causal link between these institutional constraints and misallocation is empirically difficult, largely because of the endogeneity of these institutions, but clear differences emerge between the north and south. These differences have deep historical roots and are likely a legacy of the organization of agriculture in the north before the onset of reform.

Table 14 provides a breakdown of how households acquired the land they cultivate. Most important in this context is the dominant role of land allocation by the state in the north, which at the time was heavily influenced by egalitarian considerations rather than efficiency.[19] By contrast, in the south households are much more likely to have either inherited or purchased the land they are farming. There does not appear to be differences in the role of land rental; however, other data suggest that land rental transactions in the south are much more likely to be 'arms-length' and to entail payments in cash. In the north, the contracts are primarily between relatives, and often entail no payments.[20]

Table 14: Household land acquisition (%)

Acquisition	2008			2012			2016		
	National	North	South	National	North	South	National	North	South
Allocated by state	56	64	38	55	63	37	52	60	36
Inherited	15	12	22	17	14	24	18	15	25
Purchased	8	3	20	9	4	22	10	4	22
Rented	8	7	8	6	6	6	6	6	5
Reclaimed	12	12	12	12	12	10	13	13	11
Other	1	1	1	1	0	1	1	1	1

Notes: the table reports fraction of plots by acquisition method as reported by the household. Based on balanced panel of 2,087 households.

Source: authors' own calculations based on data from VARHS.

These constraints are compounded by those that households face with respect to access to water for irrigation (Table 15), as well as restrictions on crop choice (Table 16). These issues surface in both the north and south, but are much more prominent in the north. Even as late as 2014, more than two-thirds of all households in the north reported restrictions on crop choice. Households in the north also are much more likely to report problems with respect to access to water for irrigation, as well as flooding. The pricing of water use appears to play a limited role in resource allocation in this context.

Combined, these constraints on farmers help rationalize the huge differences we observe between the north and south in the percentage of farm output that is sold. Upwards of 85 per cent of farm output is consistently sold in the south compared to one-third of so in the north (see Table 17). Indeed, much of farm output in the north is for own consumption.

Table 15: Water-related problems (%)

	Shortage of water for irrigation			Flooding			Percentage that pay for water		
	National	North	South	National	North	South	National	North	South
2008	41	51	25	25	32	13	34	35	33
2012	41	52	32	31	35	24	31	35	26
2016	47	57	33	29	40	16	27	31	19

Notes: the table reports percentage of households experiencing problems with shortage of water for irrigation or flooding and the percentage of households that are required to pay for irrigation. Based on balanced panel of 2,087 households.

Source: authors' own calculations based on data from VARHS.

[19] Estimates drawing on the 2014 VHLSS suggest much larger differences in the role of land allocation by the state between the north and south. This is largely a product of sampling, and the exclusion of the Southeast region from the VARHS data. In the south, land allocation by the state was largely limited to the South Central Coast.

[20] Sizeable differences also appear between the regions in the role of hired labour in agriculture, which may reflect a constraint as well as be a product of other constraints.

Table 16: Crop restrictions (%)

Region	2006	2008	2010	2012	2014	2016
National	58	54	39	61	57	48
North	70	64	44	73	68	52
South	42	40	33	44	43	43

Notes: fraction of households reporting government restriction on crop production for at least one plot of land. Based on balanced panel of 2,087 households.

Source: authors' own calculations based on data from VARHS.

Table 17: Fraction of crops sold (%)

Region	2006	2008	2010	2012	2014	2016
National	69	69	63	71	74	72
North	38	35	27	32	36	28
South	85	84	79	83	85	84

Notes: fraction of crops sold is calculated as the total value of sold crops divided by total reported value of produced crop. Values measured in nominal current year VND. Based on balanced panel of 2,087 households.

Source: authors' own calculations based on data from VARHS.

8 Robustness: land quality

An important concern regarding our measure of misallocation and its cost is that our measure of farmer productivity $s_{it}^{1-\gamma}$ may be capturing differences in the quality of land operated by the farm, denoted by q_{it}. To account for this potential effect, we regress our measure of farm productivity $s_{it}^{1-\gamma}$ on a set of variables related to land quality at the farm-level:

$$\ln s_{it}^{1-\gamma} = \mathbf{X}_{it}\Gamma + \Lambda_{rt} + \varepsilon_{it}, \tag{9}$$

where $\mathbf{X}_{it}$ is a vector of farm-specific variables related to the quality of the land owned by farm i; and Λ_{rt} is a region–year fixed effect for region r in year t. Specifically, we proxy land quality using the relative area of land used to grow perennials; the fraction of irrigated land (irrigation index); the fraction of land that is reported to have a flat or slight slope (flatness index); and the fraction of land that is reported as either above or below average fertility. The results from these regressions under alternative specifications are presented in Table 18.

We then calculate an adjusted measure of farm-level productivity by removing any potential influence related to differences in land quality:

$$\ln \tilde{s}_{it}^{1-\gamma} = \ln s_{it}^{1-\gamma} - \mathbf{X}_{it}\hat{\Gamma} \tag{10}$$

and an adjusted measure of effective land:

$$\ln \tilde{\ell}_{it} = \ln \ell_{it} + \frac{1}{\beta(1-\theta)\gamma}\mathbf{X}_{it}\hat{\Gamma}, \tag{11}$$

where $\hat{\Gamma}$ is the vector of estimated coefficients (listed in Table 18) and $\frac{1}{\beta(1-\theta)\gamma}$ is the inverse land production coefficient. This is included to properly account for the contribution of quality in production.

We then calculate the gains from reallocation using the adjusted measures of farm productivity $\tilde{s}_{it}^{1-\gamma}$ and land $\tilde{\ell}_{it}$ following the same procedure as before. The results are presented in Table 19, and suggest gains from reallocation similar to the baseline estimates. Adjusting for quality slightly increases misallocation in the south and slightly decreases it in the north. However, at the national and regional levels, changes in these costs over time are nearly the same as those implied by our original estimates.

Table 18: Land quality

	(1) TFP (log)	(2) TFP (log)	(3) TFP (log)	(4) TFP (log)
Rel. area of perennials	0.219***		0.251***	
	(0.0243)		(0.0246)	
Rel. value of perennials		0.106***		0.168***
		(0.0194)		(0.0195)
Irrigation index			0.248***	0.269***
			(0.0164)	(0.0166)
Flatness index			0.0117	–0.00996
			(0.0155)	(0.0152)
Above avg. fert.			0.126***	0.116***
			(0.0279)	(0.0279)
Below avg. fert.			–0.0981***	–0.0922***
			(0.0223)	(0.0223)
Observations	10,505	10,600	10,505	10,503
R^2	0.312	0.305	0.332	0.330

Notes: land-quality regressions, dependent variable in all cases is farm-level log TFP ($\log s_{it}^{1-\gamma}$). Standard errors in parentheses. $^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$.

Source: authors' own calculations based on data from VARHS.

Table 19: Productivity gain of reallocation: land-quality adjusted TFP (%)

	National		North		South	
Period	Total	Within	Total	Within	Total	Within
2006–10	62	43	64	53	44	38
2012–16	78	61	149	124	50	43

Notes: gains from reallocation using quality-adjusted measures of land and TFP. Quality-adjusted values are constructed using values in Table 18 and equations (10) and (11).

Source: authors' own calculations based on data from VARHS.

This reallocation exercise implicitly assumes that all differences in land quality are exogenous to the farmer and associated to the land. This is likely not the case and some of the differences in quality are a consequence of investment decisions by farmers, implying that they should be included in farm-level productivity. In this regard, the above reallocation exercise is an upper bound on the importance of land-quality differences for the measured gains from reallocation.

9 Conclusion

In this paper we examine important changes in agriculture in Viet Nam in the context of ongoing structural transformation in the economy. Drawing on the VARHS panel household data for the period 2006–16, we find robust growth in output averaging nearly 4 per cent per annum. This growth is a product of improvements in TFP in which the reallocation of factors across farms plays a positive role.

Nonetheless, we find that substantial misallocation of factor inputs persists across farms and that a variety of constraints facing households are preventing more rapid productivity growth. We also uncover substantive differences in the growth process and the degree of misallocation across regions in Viet

Nam: misallocation is significantly higher and rising, and productivity growth much lower in the north compared to the south.

Our analysis highlights the importance of identifying the exact features of the institutional environment that explain the differences in productivity growth and misallocation across regions and over time. Similarly, our analysis suggests the importance of changes in within-farm productivity, perhaps linked to changes in crop choices, farm size, and technology use. Investigating these dynamic linkages is an important avenue for future research.

References

Adamopoulos, T., and D. Restuccia (2014). 'The Size Distribution of Farms and International Productivity Differences'. *American Economic Review*, 104(6):1667–97.

Adamopoulos, T., L. Brandt, J. Leight, and D. Restuccia (2017). 'Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China'. Technical Report. Cambridge, MA: National Bureau of Economic Research.

Benjamin, D., and L. Brandt (2004). *Agriculture and Income Distribution in Rural Vietnam Under Economic Reforms: A Tale of Two Regions*. Washington, DC: World Bank, 2004.

Brandt, L., D. Le, G. Huong, C. Trang, G. Pham, N. Nguyen, and V. Luu (2006). 'Land Access, Land Markets, and Their Distributive Implications in Rural Vietnam'. Mimeo. Toronto: University of Toronto Department of Economics.

Do, Q.-T., and L. Iyer (2008). 'Land Titling and Rural Transition in Vietnam'. *Economic Development and Cultural Change*, 56(3): 531–79.

Foster, L., J. Haltiwanger, and C. Syverson (2008). 'Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?' *American Economic Review*, 98(1): 394–425.

Hsieh, C.-T., and P. J. Klenow (2009). 'Misallocation and Manufacturing TFP in China and India'. *The Quarterly Journal of Economics*, 124(4): 1403–48.

Markussen, T. (2017) 'Land Issues: Markets, Property Rights, and Investment'. In F. Tarp (ed.), *Growth, Structural Transformation and Rural Change in Viet Nam: A Rising Dragon on the Move*. Oxford: Oxford University Press.

Markussen, T., and F. Tarp (2014). 'Political Connections and Land-Related Investment in Rural Vietnam'. *Journal of Development Economics*, 110: 291–302.

McCaig, B., and N. Pavcnik (2013). 'Moving Out of Agriculture: Structural Change in Vietnam'. Technical Report. Cambridge, MA: National Bureau of Economic Research.

Narciso, G. (2017). 'Labour and Migration'. In F. Tarp (ed.), *Growth, Structural Transformation and Rural Change in Viet Nam: A Rising Dragon on the Move*. Oxford: Oxford University Press.

Newman, C., F. Tarp, and K. Van Den Broeck (2015). 'Property Rights and Productivity: The Case of Joint Land Titling in Vietnam'. *Land Economics*, 91(1): 91–105.

Ngai, R., C.A. Pissarides, and J. Wang (2017). 'China's Mobility Barriers and Employment Allocations'. Technical Report. London: London School of Economics and Political Science.

Restuccia, D., and R. Rogerson (2008). 'Policy Distortions and Aggregate Productivity with Heterogeneous Establishments'. *Review of Economic dynamics*, 11(4): 707–20.

Restuccia, D., and R. Rogerson (2017). 'The Causes and Costs of Misallocation'. *Journal of Economic Perspectives*, 31(3): 151–74.

Restuccia, D., and R. Santaeulalia-Llopis (2017). 'Land Misallocation and Productivity'. Technical Report. Cambridge, MA: National Bureau of Economic Research.

Tarp, F. (2017). *Growth, Structural Transformation and Rural Change in Viet Nam: A Rising Dragon on the Move*. Oxford: Oxford University Press.

World Bank (2016). *Transforming Vietnamese Agriculture: Gaining More from Less*. Washington, DC: World Bank.

Appendix: data construction

Output

Households report value and quantity of crops produced and sold in the market, allowing us to construct a crop-specific, household-level price as value divided by quantity.[21] Where possible, we use sales information to construct a market price rather than relying on the household's assessment of the crop value. We then construct a price for each crop as a weighted average of the yearly median price of the crop. We use the relative quantity of crop production in a year to construct the weight for that year.

There are two additional issues with the comparability of the data over time. First, the crop categories need to be adjusted to be consistent across years.[22] Second, the survey was substantially changed between 2006 and 2008. Notably, the questions we use to construct quantity for some crops are not available for 2006, which prevents construction of prices for some crops. We remedy this by imputing prices from the data.[23] Specifically, we regress log prices for the observed crops on fixed effects for crop, region, and year and use the predicted prices from this regression for the missing crops in 2006.[24]

Output at the household level is then constructed by aggregating quantities valued at the price index. This price index removes regional and year variations in prices, giving us a consistent measure of quantity produced across periods and regions. This gives us a comparable measure of real gross output (y in our model).

Labour

Total labour employed by the households is calculated as the sum of the household's own labour supplied to cropping activities and any labour hired outside of the household. We measure a household's own labour as the number of day equivalents supplied by the household to the production of crops.

For outside labour, we observe the amount of expenditure by the household on hired labour, but not a measure of time worked. We construct a region–year specific wage per day equivalent of work in the agricultural sector using information provided by the household on the quantity and pay members receive for working outside of the household. This wage allows us to convert expenditure on outside labour into a stock of equivalent days.

---

[21] Households are not asked to report aggregated quantities produced for some crop categories (e.g. vegetables). To remedy this, we aggregate crop quantities reported in the household's report of plot-level output by season. This includes a plot-level measure of output for the most important crop on each plot. A caveat with this approach is that households only report the quantities for the most important crop. To check the accuracy of this aggregation we compare household-level quantities for crops that are reported both at the household level and in the season's activities. The comparison shows that the measures tend to be broadly consistent, implying that most plots are only used for the production of one crop.

[22] The categories change slightly in 2008 and then again in 2012. The first change divides potatoes, cassava and sweet potatoes into three separate categories. The second change adds soybeans as a distinct category.

[23] Quantities in 2006 are missing for the categories: Vegetables, Other Annual Crops, Fruit, and Other Perennial Crops.

[24] As a check, we construct imputed quantities using the household's reported crop value and the predicted prices for crops that we observe, both values and quantities. We then compare these imputed quantities with the actual quantities reported by the household. The R-squared of regressions between the imputed and observed quantities is between 0.42 and 0.73, showing that the predicted prices capture substantial variation in farm-level prices. Note that some variation in farm-level prices is expected as our hypothesis of misallocation suggests differing prices across farmers.

Land

The main issue with calculating the quantity of land used by the household is the need to distinguish between land used for crops with land used for other non-crop uses (e.g. animal husbandry). Our measure of land is total area cultivated by the household for the production of crops.[25] Note that this includes both owned and rented-in land.

Intermediates

We observe the expenditures by the households on a variety of intermediates.[26] To convert the expenditures into a stock of intermediate inputs we construct region–year price deflators using the VHLSS. The VHLSS contains information on the prices and quantities of fertilizers used at the household level.[27] We use this to construct a median price at the region–year level and a national price common to all years. The region–year deflator is then taken to be the ratio of these two numbers. Because the VHLSS is unavailable for 2016, we adjust the 2014 deflator by the change in the national price of fertilizers from 2014 to 2016.

Capital

Our aggregate measure of the capital stock is composed of three measures of capital. First, we construct a measure of the household's non-durable capital stock using expenditure on non-durable equipment (e.g. sickles). Overall, non-durable equipment accounts for a small fraction of the aggregate capital stock (around 1 per cent).

Second, we construct a measure of durable capital using the household's reported holding of assets related to crop production.[28] For each piece of equipment, the household reports the value they believe they could obtain from selling it in the market. We deflate the values of the capital stock into common prices by using the panel dimension of the data. For example, consider a tractor owned by a household. We construct the change in price of the tractor between two surveys as the change in the reported value adjusted for depreciation.[29] For each type of equipment, we then construct an average change in prices between each pair of years and use this change to deflate the equipment values to a common price. The household's stock of durable capital is then taken to be the sum of all types of equipment valued at common prices. As a final adjustment to the durable capital owned by households, we use the panel structure of the data to fill in missing observations in the data. We use the date of purchase of the

[25] Specifically, we include all land that the household reports as being used for the production of annual or perennial crops. This excludes land used for other reasons (e.g. forestry or animal husbandry). We also exclude any land that the household reports as being left fallow for more than 48 months of the past five years.

[26] Intermediate categories: Seeds; Saplings; Chemical Fertilizers (urea, NPK, phosphate, etc.); Organic Fertilizers (self-provided); Organic Fertilizers (bought); Pesticides, herbicides; Energy, fuel (electricity, petrol, oil, lubricant, burning fuel, etc.); Minor repairs, maintenance; Payment of cultivation loan interest; Other costs (postage, advertisement, marketing, production insurance, etc.).

[27] In the VARHS, fertilizers are the most important intermediate category, accounting for around two-thirds of total nominal expenditure.

[28] Specifically, we observe the household's ownership of rice-milling machine; grain harvesting machines; pesticide sprayers; tractor; ploughs; carts.

[29] We assume that capital depreciates at a rate of 8 per cent per annum.

equipment to fill in missing observations for previous years.[30] For example, a household may report a tractor purchased in 2002 in the 2008 survey, but not in the 2006 survey.

Third, we construct a measure of capital services using the households expenditure on hired capital. We begin by constructing a price deflator for capital services using the VHLSS data. Specifically, we construct a region–year price for capital services using median earnings from capital services outside of the household.[31] We use these prices to deflate capital services to common prices, giving us a flow payment on real capital services. Next, we convert this flow measure into a stock of capital associated with capital services using the interest rate on one-year Viet Namese government bonds and a depreciation rate of 8 per cent.

[30] Note that this adjustment may introduce a downward bias in the level of capital stock reported in later years. Specifically, since we have less data to perform the adjustment in later years, we are likely underestimating the equipment in use in this year. This issue may exaggerate the downward trend in capital relative to what actually occurred. However, a comparison of the unadjusted numbers provides a similar qualitative picture as in the adjusted data. This leads us to believe that the downward trend is not being artificially created by this adjustment process. Additionally, we see the greatest change in values for earlier years, suggesting that the survey is becoming more accurate over time. Finally, we note the relatively small capital coefficient used in the analysis suggests that any errors in the capital stock will have a minor effect on the overall aggregate analysis.

[31] We observe household income from two activities that are comparable to the capital services that we observe in the VARHS: (1) ploughing and soil preparation; and (2) rice-threshing, semi-processing.