

Has the development gap between the ethnic minority and majority groups narrowed in Vietnam?: Evidence from household surveys

Tomoki Fujii 

School of Economics, Singapore Management University, Singapore

1 | INTRODUCTION

Since the beginning of *doi moi* in 1986, Vietnam has achieved truly remarkable economic development. The gross domestic product per capita on a purchasing power parity basis has soared from \$1,501 in 1990 to \$5,668 in 2015 in constant 2011 US dollars.¹ Various other development indicators have also improved significantly. Indeed, the growth story of Vietnam is often told as a success story of economic development in recent decades together with its neighbour, the People's Republic of China.

Yet, such a success story seldom conveys the whole picture. Vietnam is home to 54 officially recognised ethnic groups, where the ethnic groups of Kinh (Vietnamese), which accounts for over 85% of the population in 2009 and is the most powerful group in the government, and Ho (Chinese), which only accounts for less than 1% of the total population but traditionally has a high representation in business and commerce, are conventionally regarded as the ethnic majority groups and the remaining 52 ethnic groups as the ethnic minority groups (e.g., Baulch, Chuyen, Haughton, & Haughton, 2007; Baulch, Nguyen, Phuong, & Pham, 2010; Dang, 2014). These ethnic groups differ in culture, language, religion and historical backgrounds, although some groups are closer to each other than others. Therefore, it would not be surprising if the fruits of economic development are not equally shared across different ethnic groups.

In fact, both existing studies and this study indicate that the development gap, or inequality in development, between the ethnic minority and majority groups has not narrowed and has even widened in some aspects over the last two decades. As a villager from the Hmong ethnic minority in the Northern Mountains reportedly said, "Life is getting better but not fast enough" (*Economist* 2015). This also appears to reflect the reality that most other ethnic minority groups in Vietnam are facing.

There are at least five important reasons we should take the development gap between the ethnic minority and majority groups seriously. The first reason is that the widening development gap between the minority and majority groups is not because Vietnam's growth was anti-poor. As found by Glewwe and Dang (2011), the growth of Vietnam in 1990s was very pro-poor. However, the gap between relatively poor ethnic minority and relatively rich ethnic majority groups generally

¹World Development Indicators, 10 August 2016.

widened during the same period and continued to widen until 2014. This shows that it is not the gap between the rich and poor, but the gap between the ethnic minority and majority groups that warrants special attention.

The second reason is that geography is also not the answer to the development gap between the ethnic minority and majority groups. Because the ethnic minority group is geographically segregated from the ethnic majority group, the low welfare of ethnic minority households may stem from remoteness rather than their ethnicity per se. However, using spatially disaggregated estimates of poverty based on Vietnam Living Standards Survey (VLSS) for the year 1998² and the Vietnam Population Census for 1999, Epprecht, Müller, and Minot (2011) show that ethnicity is considerably more important than remoteness in determining household welfare.

The third reason is that the development gap may undermine social cohesion and lead to violence and civil unrest. For example, the Central Highlands witnessed major protests against the central government involving hundreds or thousands of ethnic Hmong in 2001, 2004 and 2011, which resulted in the use of force. While land disputes and lack of religious freedom and autonomy are generally considered the direct causes of these protests, many have pointed out persistent poverty as an important contributing factor (see, e.g., Writenet, 2006). The Vietnam News (2006), a state-run news agency, also argues that most of the local hill-tribe ethnic people in the Central Highlands are still trapped in poverty and thus prone to being deceived by hostile forces to protest against the government and state.

The fourth reason is ethical. The development gap between the ethnic minority and majority groups is a cause for concern, because it may reflect inequality of opportunity. In the growing body of empirical literature on inequality of opportunity (Aaberge, Mogstad, & Peragine, 2011; Bourguignon, Ferreira, & Menédez, 2007; Checchi & Peragine, 2010; Ferreira & Gignoux, 2011; Hassine, 2011), the observable factors or variables that affect the outcome of interest (e.g., income or consumption) are often categorised into “circumstances” and “efforts” following Roemer (1998), where the former (latter) are beyond (within) the scope of individuals’ responsibility. While the details of measurement vary across studies (see Pignataro, 2012 and Ramos and Van de Gaer, 2016 for recent relevant surveys), the inequality of opportunity is typically defined as the inequality due to the circumstances and thus considered to be more ethically unacceptable than the inequality of outcome. Inequality of opportunity is also important from the growth perspective because higher inequality of opportunity is associated with lower economic growth (Marrero & Rodríguez, 2013; Marrero, Rodríguez, & van der Weide, 2016). Because the individual’s ethnicity is clearly a circumstance variable, the development gap between the ethnic minority and majority groups is important from the perspectives of ethics and economic growth.

Finally, if the development gap stems from a discrimination against the minority group, eliminating the gap is clearly an important objective on its own. Under the Constitution of the Socialist Republic of Vietnam, all the ethnicities are equal and all acts of discrimination against and division of the ethnicities are prohibited. Minorities have the right to speak their own languages and to preserve and develop the “positive” elements of their cultures. However, cultural traits that are deemed negative by the Kinh-led state are not allowed and extreme measures were sometimes taken to eradicate such traits in the past (Rambo, 2003). Moreover, some ethnic minority groups may also have been targeted for repression partly because they fought with American and South Vietnamese troops in the Vietnam War (*Economist*, 2015).

²To be exact, the data collection took place between December 1997 and December 1998. However, because the majority of households were surveyed in 1998, we shall simply take the reference year to be 1998 and call this data set VLSS 1998. Likewise, an earlier round of VLSS, which took place between October 1992 and ended in October 1993, will be referred to as VLSS 1993 below.

Even without overt ethnic discrimination, the development gap is still an important cause for concern if it reflects inadequate or inappropriate policies. In particular, when the policies do not take into account the reality of the lives of ethnic minority groups, they may easily fail even if they are well intended. There are indeed a number of cases where policies aimed to narrow the development gap yielded only disappointing results. For example, piglets and rice seeds are distributed to mountaintop farms, but the breeds and varieties are best adapted to the hotter lowlands. School textbooks are mainly published in Vietnamese rather than in local languages (*Economist*, 2015). Efforts to develop standardised systems of writing for minority languages and to promote their use in primary education and anti-illiteracy campaigns only yielded disappointing results because existing local writing systems were ignored (Rambo, 2003). Recent state forest and land allocation policies also made the traditional lives of some ethnic minority groups more difficult. For example, three quarters of Sa Pa District in Lao Chai Province is located within Hoang Lien National Park, which has been designated a protected forest since 2002. Collecting plants and animals, lighting fires and grazing animals there are prohibited, and this reduced the access to historically important sources of cash income for Hmong households living there (Tugault-Laeur & Turner, 2011). Various other types of mismatches between what is needed and what is provided are also reported in World Bank (2009).

The primary purpose of this paper was to advance the understanding of the ethnicity-based development gap in Vietnam. We start by reviewing the existing literature in Section 2. We then document the changes in various development indicators for the ethnic minority and majority groups between 1993 and 2014 in Section 3. We show that the standard of living has improved for both ethnic minority and majority groups, but there is no indication that the gap between them clearly narrowed over our study period except for the use of electricity as a main source of lighting.

In Section 4, we investigate the returns to characteristics. Our analysis shows that there is a persistent structural gap between the ethnic minority and majority groups. However, the nature of the structural difference between these two groups also changed over time. The poverty decomposition analysis presented in Section 5 indicates that the structural change led to an increase in the development gap between the ethnic minority and majority groups until 2006. However, there are indications that the structural change actually helped to narrow the gap between them since then.

Because the poverty decomposition analysis focuses on those who are under the poverty line, it is insensitive to the upper tails of consumption distribution. To see whether relaxing our focus on the lower tail of distribution qualitatively alters our empirical results, we compare our poverty decomposition analysis with inequality decomposition used in the inequality of opportunity literature in Section 6. We find that the results of poverty and inequality decompositions are broadly consistent with each other.

One possible reason for the persistent development gap is the mistargeting or leakage of aid such that aid resources allocated to the ethnic minority group are captured by local elites or ethnic majority households. To see whether this is the case, we analyse the incidence of aid in Section 7. We find that this is not the case and that existing assistance actually disproportionately goes to the ethnic minority group. Finally, we provide some discussion in Section 8. In particular, we argue that adopting more minority-appropriate policies would be necessary for lifting poor ethnic minority households out of poverty and narrowing the development gap between the ethnic minority and majority groups.

2 | PREVIOUS STUDIES ON THE DEVELOPMENT GAP

Before the beginning of this century, there was only a handful of microeconomic studies on or related to the development gap between ethnic minority and majority groups in Vietnam partly

because of the lack of reliable and relevant data. To our knowledge, one of the earliest published studies is Houghton and Houghton (1997), who show that children born into the ethnic minorities are significantly more likely to be stunted, using VLSS 1993. Their estimate indicates that such children tend to have a lower height-for-age Z-score than their ethnic majority counterparts by 0.2. This effect is large, but they also acknowledge that their estimate may be picking up a geographic effect, since the minority peoples are disproportionately found in more remote and mountainous areas.

van de Walle and Gunewardena (2001) also use VLSS 1993 to analyse the sources of inequality between the ethnic minority and majority groups. They use a subsample of rural households in what they loosely call northern Vietnam, which includes northern regions and the Central Highlands. They then regress the logarithmic consumption expenditure per capita on a variety of covariates, including household demographic characteristics, education and different types of land areas cultivated by the household, to estimate the demographic effects and returns to education and land for the ethnic minority and majority groups. Using these estimates, they carry out Oaxaca–Blinder decompositions (Blinder, 1973; Oaxaca, 1973) to explain the difference in mean logarithmic consumption between the two groups.

van de Walle and Gunewardena (2001) find, among others, that the estimated demographic effects and returns to education and land are different between ethnic minority and majority groups and the difference is statistically significant jointly. Therefore, their decomposition results suggest that it is appropriate to use a separate regression model for ethnic minority and majority groups. Their result also suggests that commune-level characteristics may be important. While their decomposition results are sensitive to the choice of sample (i.e., whether to restrict to mixed communes where both minority and majority groups reside), they find that a sizable fraction of the consumption gap between the ethnic minority and majority groups can be attributed to the different returns to characteristics such as demographics, education and land. This study also investigates whether and by how much the returns to characteristics are different between the ethnic minority and majority groups and how this difference has evolved over time using more recent data.

Using VLSS 1993 and 1998, Litchfield and Justino (2004) show that the ethnic minority groups are poorer than the ethnic majority groups and the difference in the poverty rates between the ethnic minority and majority groups widened between 1993 and 1998. Similarly, Baulch et al. (2007) document the development gap between the ethnic minority and majority groups using the same set of data. Despite the overall progress of economic development in Vietnam during this period, the standard of living of minorities in the Central Highlands and, to a lesser extent, the Northern Highlands lagged behind those of the ethnic majority groups. Analysing the 3% enumeration sample of the 1999 census, Baulch et al. (2007) also show that the enrolment rates of ethnic minority groups at the lower secondary level tend to be lower than that of the ethnic majority group.

Further, Baulch et al. (2007) conduct a decomposition analysis similar to van de Walle and Gunewardena (2001) using VLSS 1998. As with van de Walle and Gunewardena (2001), they find that ethnic minority and majority groups are structurally different. Their results indicate that more than half of the gap between the ethnic minority and majority groups can be attributed to the different returns to characteristics in most specifications. This means that there will remain a sizable gap between the ethnic minority and majority groups, even if the distribution of characteristics between the two groups is hypothetically made equal.

A similar implication can be also drawn from Nguyen, Albrecht, Vroman, and Westbrook (2007), who also use VLSS 1993 and VLSS 1998. While their focus is on the disparity between urban and rural areas, their regression analysis of household consumption expenditure per capita shows that there is a penalty for the ethnic minority status. Furthermore, they find that this penalty increased between 1993 and 1998.

Baulch et al. (2010) provide a piece of evidence that the development gap between the ethnic minority and majority groups in the 1990s persisted until this century. Using the Vietnam Household Living Standard Survey (VHLSS) for 2002, 2004 and 2006 as well as VLSS 1998, they find that the difference in the consumption level between the ethnic minority and majority groups widened between 1998 and 2006. They also report worrying deterioration in the nutrition indicators of children between 2 and 4 years of age. In 1998, the proportion of stunted children (i.e., children whose height-for-age is more than two standard deviations below the median of the reference population) in this age group increased between 1998 and 2006 for rural ethnic minorities from 54% to 57%. During the same period, the corresponding proportion for the ethnic majority group has dropped from 53% to 37%. Widening gap between the ethnic minority and majority groups was also observed for the prevalence of stunting for children under 2 years of age.

In addition to simple descriptive statistics, Baulch et al. (2010) also present the results of Oaxaca–Blinder decomposition for the years 1998, 2004 and 2006. Unlike the previous studies by van de Walle and Gunewardena (2001) and Baulch et al. (2007), they conduct a Oaxaca–Blinder decomposition analysis based on quantile regressions using the technique first proposed by Machado and Mata (2005). Regardless of the method used and the study year, the results are broadly consistent with previous studies. That is, a sizable fraction of the observed difference in per capita expenditure between ethnic minority and majority groups can be attributed to the difference in returns to observed characteristics.

The decomposition results presented by van de Walle and Gunewardena (2001), Baulch et al. (2007, 2010) all indicate that the returns to characteristics are an important driver of the development gap between the ethnic minority and majority groups. Therefore, as Baulch et al. (2007), argue, if policymakers want to close the gap between minority and majority living standards, while maintaining ethnic identities, then it will not be sufficient simply to improve minority education or provide minority households with more land. This also means that antipoverty programmes that are geared towards the minority will have to look different from those geared towards the majority.

Dang (2014) also describes the development gap between the ethnic minority and majority groups in Vietnam using VLSS 1993 and 1998 as well as VHLSS 2002, 2004, and 2006. According to his calculations, the poverty rate for the ethnic majority group dropped by 71% from 54% to 10%, whereas the corresponding figure for the ethnic minority group only dropped by 42% between 1993 and 2006. The contrast is even more striking in the rural area. The ethnic majority group reduced the poverty rate from 62% to 14%, whereas the ethnic minority group reduced poverty rate from 88% only to 54% during the same period. He also finds that the ethnic minority group lags behind the ethnic majority group in a variety of indicators such as income, education, health, prevalence of child labour and access to community services despite various efforts made by the government to close the gap.

Among the studies mentioned above, this study particularly closely relates to van de Walle and Gunewardena (2001), Baulch et al. (2007, 2010), and Dang (2014) because our focus is also on the gap between the ethnic minority and majority groups. Further, the specification of the regression model is similar to those used in these studies.

However, this study differs at least in three important aspects. First, in addition to the VLSS and VHLSS data used in previous studies (i.e., VLSS 1993 and 1998 and VHLSS 2002, 2004 and 2006), we use more recent rounds of VHLSS data for the years 2008, 2010, 2012 and 2014. This allows us to provide more up-to-date analysis and identify the persistent development gap more clearly.

Second, we closely examine various non-consumption development indicators in addition to consumption expenditure per capita, which is often used as the main and sometimes only outcome of interest in existing studies. We do this because the consumption measure may be

sensitive to changes in the format of questionnaires (Lanjouw & Lanjouw, 2001) as the great Indian poverty debate clearly exemplifies (Deaton & Kozel, 2005). This point is particularly relevant to our application because there were a few important differences in the design of the consumption module between VLSS and VHLSS series. That is, the number of food (and drink) items included in VLSS 1998 is 45 as opposed to 58 in VHLSS 2002. Further, the recollection period for non-food items has changed from the last 4 weeks in VLSS 1998 to the last 12 months in VHLSS 2002. The recollection period for some regularly consumed items was further changed to 1 month from VHLSS 2010. There are also a few minor modifications made over time. These issues may potentially make the direct comparison of the consumption measure across surveys problematic. This is particularly true for the comparison between the VLSS and VHLSS series.

Therefore, it is important to compare the trends in development gaps using multiple outcome variables. In this sense, our paper is also related to Baulch and Masset (2003), who compare the transition of consumption poverty against the transition of non-consumption poverty (malnutrition or being out of school) using VLSS 1993 and 1998. Their study finds that consumption poverty is less persistent than non-consumption poverty. They interpret it as a reflection of irreversibilities in education and nutrition and underscore the importance of taking into account the multidimensionality of poverty.

Unlike Baulch and Masset (2003), however, the purpose of comparing multiple outcomes in this study is to assess whether the results based on potentially incomparable consumption data are plausible. Hence, we choose to use some measures of asset holdings and housing conditions partly because of the data availability (e.g., nutrition measure is not available in some VHLSS rounds) but mostly because these indicators are expected to directly reflect the standard of living of the household. We indeed find that the trend of the consumption gap between the ethnic minority and majority groups broadly agrees with the corresponding trend of the nonconsumption gap.

Finally, we apply a new poverty-focused decomposition method proposed by Fujii (2015) to our data instead of the popular Oaxaca–Blinder decomposition, which does not distinguish between those below and above the poverty line. As elaborated later, our method has several distinct advantages over the Oaxaca–Blinder decomposition when we are primarily interested in development gaps.

3 | DEVELOPMENT GAP BETWEEN 1993 AND 2014

This study uses a total of nine rounds of VLSS and VHLSS surveys between 1993 and 2014. Both VLSS and VHLSS data sets are partial panel data. Following van de Walle and Gunewardena (2001), we only use rural households in what is loosely called northern regions. Further details of the data are found in Appendix A.

Table 1 describes the means of some key characteristics of the data broken down by the ethnic majority (Kinh and Hoa) and minority (all other ethnicities) groups for the years 1993, 2002, 2008 and 2014. While there is a large heterogeneity across ethnic groups even within the ethnic minority group, we aggregate them because the number of observations for each ethnic group is too small to conduct a meaningful analysis.³

³This limitation applies to virtually all studies using VLSS and VHLSS data. United Nations Development Programme (2013) reports poverty and income for 13 ethnic groups including “others” using a data set focused on ethnic minority households.

TABLE 1 Means of key variables for the ethnic minority and majority groups for the years 1993, 2002, 2008 and 2014 weighted by the population expansion factor

Year Minority/Majority	1993		2002		2008		2014	
	Min	Maj	Min	Maj	Min	Maj	Min	Maj
Household's demographic characteristics								
Female-headed household (%)	10.9	18.6	7.0	14.5	6.5	16.9	9.2	17.5
Age of household head	40.6	43.7	43.3	46.3	43.4	49.3	43.4	51.1
Head is married (%)	90.9	88.7	93.4	89.6	94.0	87.4	92.3	86.2
Household size	6.1	5.4	6.1	4.8	5.7	4.5	5.1	4.3
Ratio of children aged 14 and below (%)	43.4	39.3	37.0	29.8	30.6	23.0	29.9	23.0
Ratio of elderly aged 65 and above (%)	5.0	5.7	5.0	7.4	4.8	9.2	4.3	9.5
Household head's education level								
No education or primary incomplete (%)	40.5	21.7	47.9	16.0	41.7	11.2	41.3	10.7
Primary completed (%)	25.4	23.9	27.9	21.1	29.8	22.9	29.6	21.9
Lower secondary completed (%)	21.0	35.3	17.5	44.8	20.6	47.3	20.8	45.3
Upper secondary completed (%)	11.8	17.8	6.2	16.3	7.4	16.3	6.8	18.7
University or higher (%)	1.3	1.3	0.6	1.9	0.4	2.2	1.5	3.4
Household head's sector of employment								
Head is not working (%)	5.5	6.6	4.5	8.5	3.3	10.1	2.9	12.0
Working in primary sector ^a (%)	90.7	78.2	88.0	60.5	87.2	55.0	84.6	42.9
Working in secondary sector ^a (%)	1.5	7.3	1.7	16.5	3.7	26.9	6.8	25.9
Working in tertiary sector ^a (%)	2.3	7.8	5.8	14.6	5.7	7.9	5.7	19.2
Land area cultivated by the household								
Area for rice (m ²)	4,908	4,641	5,484	3,671	8,337	6,170	4,053	2,848
Area for other annual food crops (m ²)	4,252	1,328	4,602	983	6,084	909	4,167	724
Area for annual industrial crops (m ²)	731	559	688	425	708	401	467	413
Area for perennial industrial crops (m ²)	270	221	907	458	622	863	884	727
Area for fruit trees (m ²)	29	55	224	158	173	73	95	74
Consumption and poverty								
Per capita daily consumption ^b	1.3	1.6	1.6	2.2	2.5	3.8	3.8	7.0
\$1.90-a-day poverty (%)	88.6	76.1	77.6	52.3	35.9	9.8	13.8	1.0

(Continues)

TABLE 1 (Continued)

Year Minority/Majority	1993		2002		2008		2014	
	Min	Maj	Min	Maj	Min	Maj	Min	Maj
\$3.10-a-day poverty (%)	98.6	95.1	95.0	84.8	77.9	46.2	49.9	6.9
Number of observations	386	1,917	3,145	10,214	1,090	3,068	1,187	2,761

Notes: ^aPrimary sector includes agriculture, fishing, hunting and mining. Secondary sector includes manufacturing sector. Tertiary sector includes services.

^bIn 2011, International US Dollars converted from nominal Vietnamese Dong using the purchasing power parity (PPP) conversion factor for the gross domestic product of 1,627, 3,033, and 4,838, and 7,682 for the years 1993, 2002, 2008 and 2014, respectively.

The top part of Table 1 describes the demographic characteristics of the minority and majority households. It shows that the proportion of individuals living in a female-headed household for the minority households is lower than that for the majority households. Whether they are in the ethnic minority or majority group, most household heads are married. Table 1 also shows that ethnic minority households have a larger number of members, younger head, higher ratio of children, and a lower ratio of elderly people than ethnic majority households.

For the ethnic majority group, the average educational attainment of household head has clearly increased between 1993 and 2014. On the other hand, the average educational attainment of household head for the minority group slightly dropped during the same period. As we show in Figure 3a below, the gap between the ethnic minority and majority groups in the education level of household head widened between 1993 and 2014.

The sector of employment, where employment hereafter includes self-employment, is also different between the ethnic minority and majority groups. As of the year 1993, most household heads in both minority and majority households in rural northern Vietnam were employed in the primary sector with rice farming being the most important activity. While both groups reduced the share of heads employed in the primary sector between 1993 and 2014, the reduction was far larger for the ethnic majority group. Accordingly, the secondary and tertiary sectors have substantially increased their importance for the employment of ethnic majority household heads between 1993 and 2014. However, the shares of secondary and tertiary sectors remained low for the ethnic minority group.

Minority and majority households also differ in terms of land areas they cultivate. We categorise the land cultivated by each household into the following five categories: (i) rice; (ii) other annual food crops (such as tubers, roots and leafy vegetables); (iii) annual industrial crops (such as soybean, peanut, sesame, sugar cane, tobacco, cotton, jute and rush); (iv) perennial industrial crops (such as tea, coffee, rubber, black pepper, coconut, mulberry and cashew); and (v) fruit trees. In 2014, the ethnic minority groups have on average larger cultivated areas than ethnic majority households in each of these categories. In particular, the areas cultivated for food crops other than rice for the ethnic minority group are much larger than those for the ethnic majority group. It should be noted, however, that the average is taken over all households, which include non-farmers and part-time farmers. Further, the quality of land may be different between the minority and majority households. Therefore, a larger cultivated area does not necessarily translate into a higher yield and a higher standard of living for the household. Table 1 also shows that the way the pattern of cultivation changed over time is different between the minority and majority households. In particular, there was a significant shift away from rice production for the ethnic majority group between 1993 and 2014. However, the corresponding change for the ethnic minority group was modest.

The bottom part of Table 1 reports daily living consumption expenditure per capita. We take the nominal consumption measure provided in the data for the years 1993–2008 and construct a

comparable measure for the years 2010–2014. These measures are then converted to 2011 US dollars using the purchasing power parity conversion factor obtained from the World Development Indicators. As noted earlier, the consumption module of the survey differs across different survey rounds and thus strict comparability of the consumption measure, particularly between the VLSS and VHLSS series, is debatable at best.

However, the consumption ratio between the ethnic minority and majority groups is likely to be robust to survey design unless ethnicity-specific consumption items are included in some rounds but not others. Therefore, ignoring the potential comparability issue, Table 1 shows that the consumption gap has substantially widened over the last two decades. The ratio of average per capita consumption between minority and majority households dropped from 81% (= 1.3/1.6) in 1993 to 54% (= 3.8/7.0) in 2014.

We also computed the ratio of individuals whose daily consumption per capita falls below the international poverty lines of \$1.90-a-day and \$3.10-a-day in 2011 US dollars. As the table shows, most people were poor in 1993. While both the majority and minority groups reduced poverty significantly over the last two decades, the reduction was much faster for the former. In particular, about half of the people in the ethnic minority group are still under the \$3.10-a-day poverty line, whereas only less than 7% of the ethnic majority group falls below this poverty line, showing that the poor in the minority group lagged behind the poor in the majority group.

It is of interest to see whether the consumption gap between the ethnic minority and majority groups has widened for the poor or the rich. Therefore, we plot the difference in the logarithmic consumption per capita at different consumption quantiles between the ethnic minority and majority groups from 1993 to 2014 in Figure 1, which shows that the middle part of the graph is almost horizontal for each year. This means that the consumption gap is fairly similar across different quantiles, except for the very top and very bottom, in a given year. For example, the consumption for the majority group was about 20% larger than that for the minority group in 1993 for any quantile between 20th and 80th percentiles of consumption distribution. This gap had increased to about

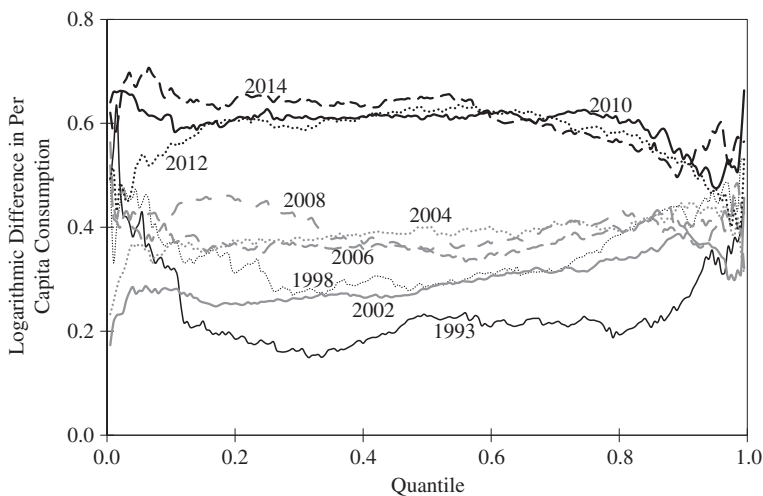


FIGURE 1 Logarithmic difference in household consumption per capita between the ethnic majority and minority groups at each consumption quantiles. The thin lines are for VLSS 1993 (solid) and 1998 (dot). The grey thick lines are for VHLSS 2002 (solid), 2004 (dot), 2006 (dash), and 2008 (dash-dot). The black thick lines are for VHLSS 2010 (solid), 2012 (dot) and 2014 (dash)

60% by 2010 and stayed around that level until 2014. A relatively large increase in consumption gap occurred between 1993 and 1998, between 2002 and 2004, and between 2008 and 2010.

An alternative way to look at the development gap is to see the proportion of ethnic composition at each consumption quantile. Figure 2 shows the results of local-mean smoothing regressions of the indicator for the individuals residing in a household whose head belongs to the ethnic minority group on the consumption quantile for different years. The area below each regression line represents the share of minority individuals in northern rural Vietnam in a given year. This share has increased from 19% to 27% between 1993 and 2014 because of the differential migration, mortality and fertility rates between the ethnic minority and majority groups. Therefore, it is not surprising that the regression line generally shifted upwards over time.

Figure 2 also shows much of the upward shift occurred for lower consumption quantiles. In 1993, about 40% of the poorest of the poor were accounted for by the ethnic minority group. This share has increased to about 80% by 2014. This does not mean that ethnic minority individuals in the lowest consumption quantiles became poorer because the actual consumption level at each quantile has increased over time thanks to the rapid economic growth during our study period. However, the graph does show that the ethnic majority group generally shifted upwards in the consumption distribution whereas the opposite is true for the ethnic minority group.

It should be reiterated here that the consumption expenditure is not directly comparable between VLSS and VHLSS because of the difference in the consumption module. While we do not see much change in the pattern of consumption gaps between 1998 and 2002, we cannot exclude the possibility that the change in the design of the consumption module may have affected the results shown above. Therefore, it is also useful to see whether the widening gap is also observed from other development indicators.

Figure 3a plots the proportion of individuals living in a household headed by a person with at least lower secondary education for the ethnic minority and majority groups. This proportion has

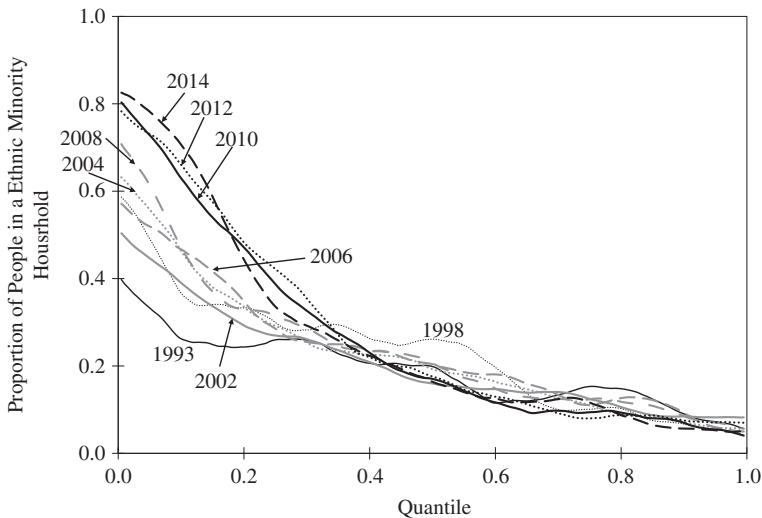


FIGURE 2 Estimates of the proportion of ethnic minority individuals obtained by local-mean smoothing regression of the indicator for the individuals residing in a household whose head belongs to the ethnic minority group on the consumption quantile for different years with the Epanechnikov kernel and the bandwidth of 0.04. The thin lines are for VLSS 1993 (solid) and 1998 (dot). The grey thick lines are for VHLSS 2002 (solid), 2004 (dot), 2006 (dash) and 2008 (dash-dot). The black thick lines are for VHLSS 2010 (solid), 2012 (dot) and 2014 (dash)

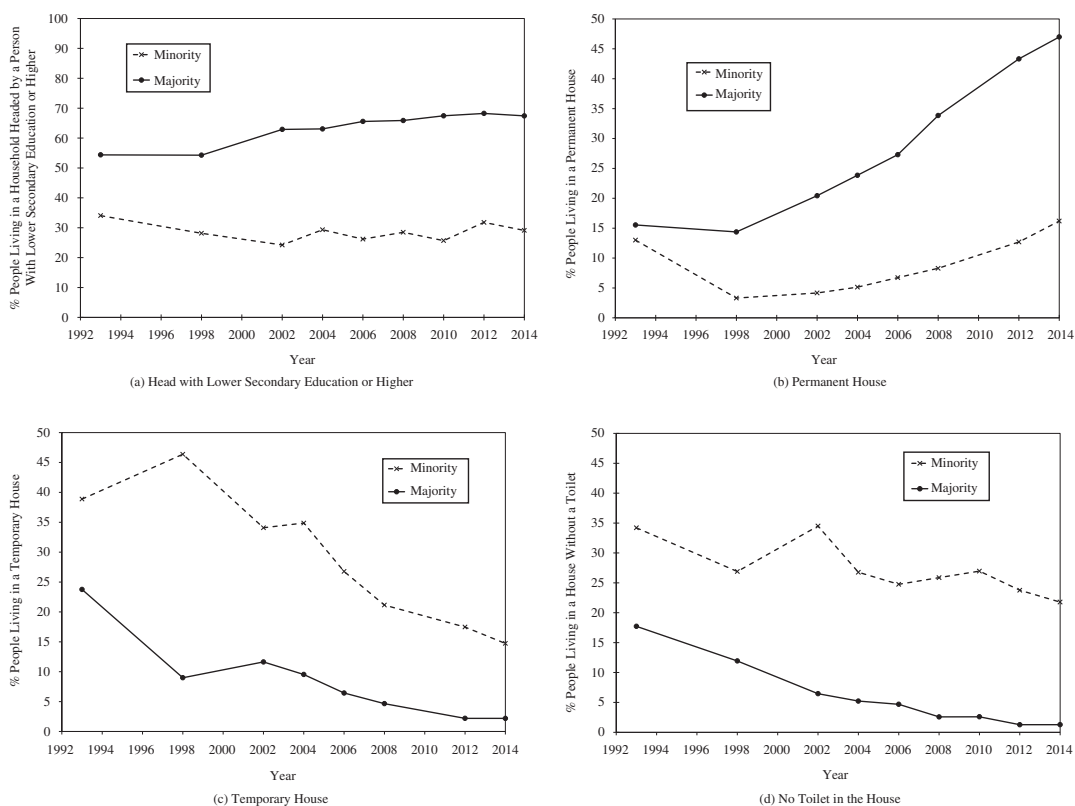


FIGURE 3 Education and housing indicators for the ethnic minority and majority groups between 1993 and 2014

increased for the ethnic majority group, whereas it has declined slightly for the ethnic minority group between 1993 and 2014. Therefore, the difference in the potential earning capacity of households appears to have increased.

Figure 3b and c plots the proportions of individuals living in permanent and temporary houses, respectively.⁴ Permanent houses include villas and multistory houses and are made with strong materials such as concrete and bricks. On the other hand, temporary houses are makeshift houses typically made with weak materials such as bamboo, straw and leaves. As these figures show, the share of permanent houses increased and the share of temporary houses decreased for both the ethnic minority and majority groups. However, the pace of improvement was significantly different between the ethnic minority and majority groups. The proportion of people living in a permanent house was similar between the ethnic minority and majority groups in 1993. However, the proportion for the ethnic majority group became far larger than that for the minority group by 2014. The difference in the share of people living in a temporary house between the two groups did not narrow much between 1993 and 2014.

Figure 3d shows that the proportion of people living in a house without a toilet on its premises declined for both groups. However, the rate of decline is faster for the majority group than for the minority group. We again see a persistent gap between the ethnic minority and majority groups.

⁴These variables were not included in VHLSS 2010.

Figure 4a plots the proportion of individuals who obtain water for cooking and drinking primarily from a private tap. Almost no one was getting water from a private tap in 1993 in our sample. While the proportion remains low, it has steadily increased for the ethnic majority. On the other hand, this proportion is still only around 1% for the ethnic minority group in 2014.

Figure 4b plots the proportion of individuals whose main source of lighting is electricity. It shows that electricity has spread rapidly during our study period. Only 57% of the ethnic majority individuals used electricity as the main source of lighting in 1993. Since 2006, however, almost 100% of the individuals in an ethnic majority household use electricity as a main source of lighting. The pattern is somewhat similar for the ethnic minority households. Almost no one used electricity as the main source of lighting in 1993. Now the proportion is as high as 80%. While the gap between the ethnic minority and majority groups in the use of electricity as the main lighting source narrowed until 2008, the gap stayed at a similar level after then.

Figure 4c and d plots the fraction of individuals living in a household with a motorcycle and television, respectively. As these figures show, both motorcycle and television have spread rapidly over the last two decades for both the ethnic minority and majority households. The minority–majority gap in the share of individuals with these assets in the household has increased, although the increase is only marginal for motorcycle.

The results presented in this section confirm and reinforce the findings of previous studies discussed in Section 2. In 1993, there was already a development gap between the ethnic minority and majority groups. While a variety of development indicators have clearly improved since 1993

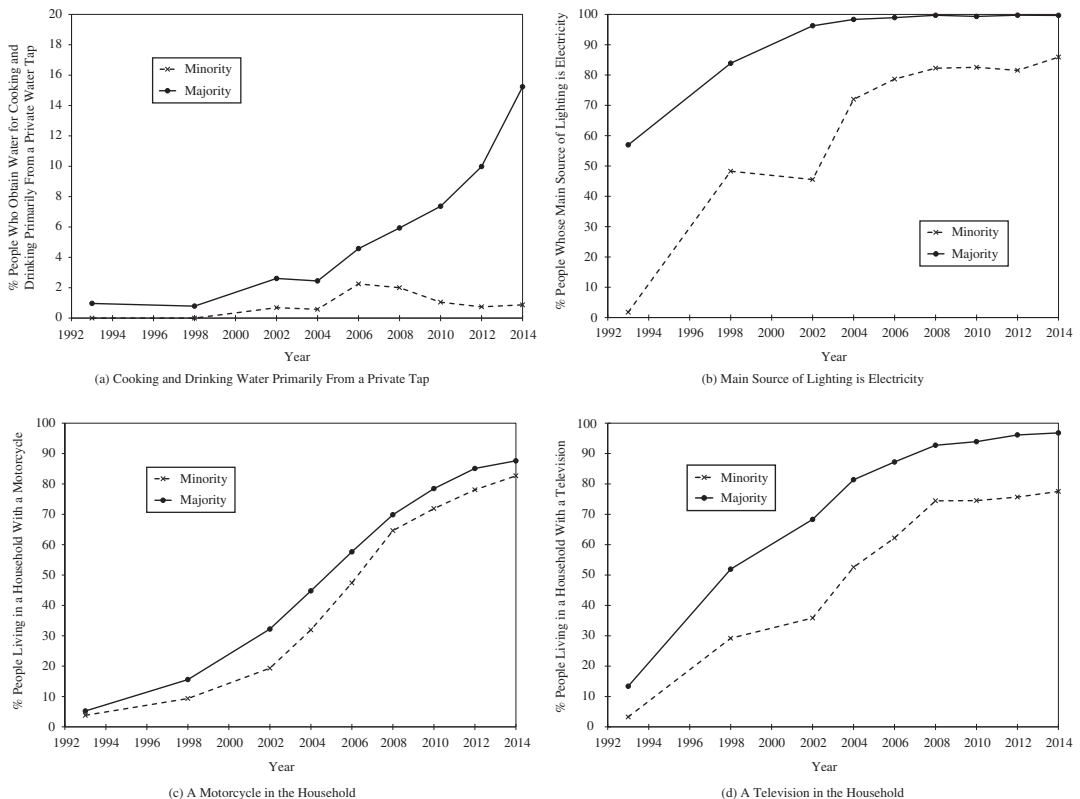


FIGURE 4 Utility and asset indicators for the ethnic minority and majority groups between 1993 and 2014

for both groups, the gap between them remained or even widened for all the indicators we considered except for the use of electricity as the main lighting source.

It should be noted here that our results are not driven by selective migration of rich minority households to the urban areas or southern Vietnam. While the results presented in this section are based on a subsample of households in northern rural Vietnam, the widening gap is also observed when the sample for the entire country is used.

The preceding discussion only provides a snapshot of development gap between the ethnic minority and majority groups. It is, however, interesting to see whether the transition of poverty is also different between the minority and majority groups. To this end, we use six sets of balanced panel data for years 1993–1998, 2002–2004, 2004–2006, 2006–2008, 2010–2012 and 2012–2014. Because the VHLSS data are a rotating panel data set, the set of households for the year 2004 is different between the panel VHLSS 2002–2004 and VHLSS 2004–2006. A similar remark applies to the years 2006 and 2012. It should be reiterated that there are no panel households between 1998 and 2002 and between 2008 and 2010. These six panel data sets are also used in the next three sections.

Figure 5a provides the proportion of people who got out of poverty among those who were initially poor under the \$1.90-a-day poverty line. The figure shows, for example, that about one-half of the ethnic majority people who were poor in 1993 were out of poverty in 1998. The corresponding proportion for the ethnic minority is roughly half as much. The figure therefore shows that the probability of getting out of poverty for the ethnic minority group was consistently below that for the ethnic majority group during all the six time periods we studied. Figure 5b provides the proportion of people who fell into poverty among those who were not poor initially. This figure shows that the ethnic minority group is more vulnerable to poverty than the ethnic majority group. These observations indicate that the ethnic minority group has been consistently more likely to be trapped in poverty than the ethnic majority group.

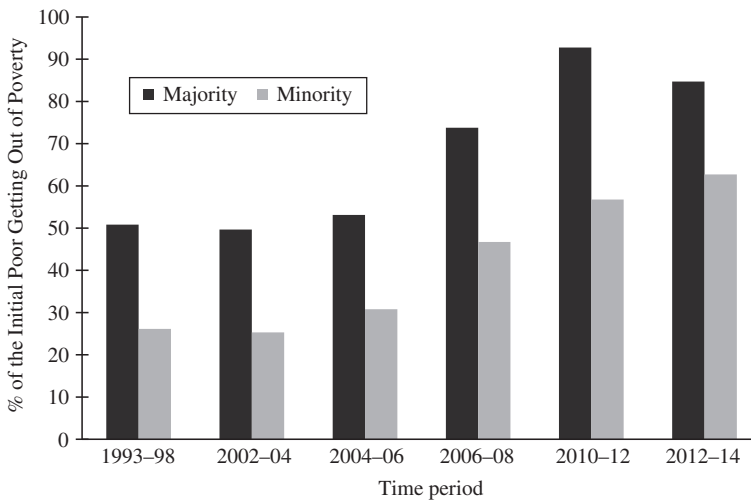
4 | RETURNS TO CHARACTERISTICS

One reason why there is a development gap between ethnic minority and majority groups is that the returns to characteristics of the households, such as education, sector of employment and cultivated land areas, are different. In this case, even when both groups have an identical distribution of characteristics, the gap between the ethnic minority and majority groups will still remain. To see whether this is the case, we run a series of regressions with the following specification:

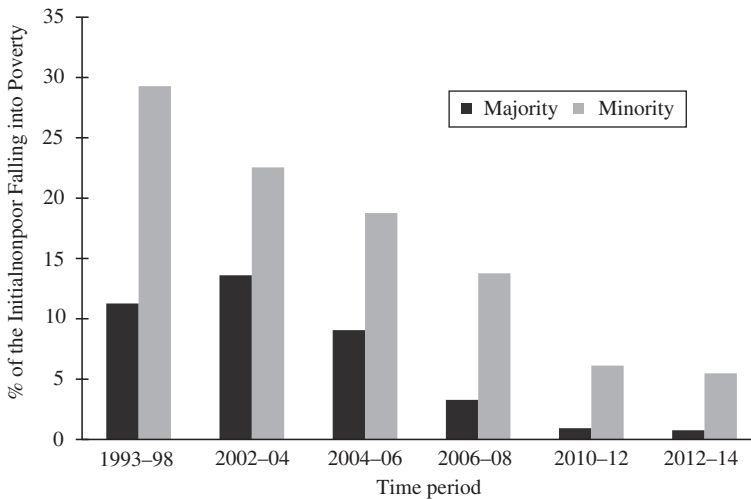
$$\ln y_{ht}^e = x_{ht}^e \beta_t^e + v_{p(h)t}^e + \epsilon_{ht}^e, \quad (1)$$

where the superscript e denotes the ethnicity (0 for minority and 1 for majority) and the subscripts h and t denote household and time, respectively. The variables y_{ht}^e and x_{ht}^e are the household consumption per capita and a row vector of covariates, respectively, and β_t^e is a column vector of regression coefficients. The error term has two components. The first term $v_{p(h)}$ is the provincial effect, where $p(\cdot)$ is a mapping from the household to the province it is located in. The second term ϵ_{ht}^e is the idiosyncratic error term that is independently and identically distributed.

Existing regression studies reviewed in Section 2 analyse only one time point at a time. That is, they compare between β_t^0 and β_t^1 for a particular time period. Unlike these studies, we also analyse the difference in the change of coefficients over time between the ethnic minority and majority groups with panel data sets so that the set of households is fixed for a given period of analysis.



(a) Proportion of People who Got Out of Poverty Among Those Who Were Poor Initially



(b) Proportion of People Who Fell Into Poverty Among Those Who Were Not Poor Initially

FIGURE 5 Poverty transition for ethnic majority and minority groups under \$1.90-a-day poverty line

The choice of the specification above and the use of panel data sets are driven by two considerations. First, the regression results based on Equation (1) using panel data sets directly feed into our poverty decomposition analysis presented in the next section. While it is also possible to run similar panel regressions with household fixed effects, the regression coefficients (in a typical specification) are fixed over time for a given ethnic group and thus no structural change is allowed, a feature unattractive for our purpose.

Second, our choice also has an advantage that the potential presence of confounding factors for the ethnicity would not be as important as it would in a typical cross-sectional analysis. Because we use the same set of households between the initial and terminal time periods, the omitted-variable bias in the coefficient estimate due to the correlation between unobserved characteristics and the ethnicity of the household head (e.g., remoteness of the location of residence) is likely to



cancel out, to some extent, by taking the difference in the coefficients between the initial and terminal periods. Therefore, the difference between the ethnic minority and majority groups in the change in the regression coefficients between the two time periods, or the difference-in-differences (DiD) in coefficients, is less subject to the omitted-variable bias.

For each year in each panel data set and for each of the ethnic minority and majority groups, we run a regression of the logarithmic consumption per capita on a rich set of covariates. Table 2 presents regression results for the years 1993 and 2002 and Table 3 for the years 2008 and 2014 out of a total of 24 regressions.⁵

As with previous studies such as Baulch et al. (2007) and van de Walle and Gunewardena (2001), we include household's demographic and educational characteristics as well as the five types of cultivated land areas discussed in the previous section. However, the exact set of covariates used in this study differs from existing studies. This is because the format of relevant questions changed over years and thus we needed to transform some variables to maintain comparability across years.

We also include the indicator variables for the household head's sector of employment, where the base category is non-working household head. Inclusion of the head's employment sector is potentially important because the importance of secondary and tertiary sectors has substantially increased in Vietnam over the last two decades. Therefore, the differences in employment pattern between the ethnic minority and majority groups may explain the development gap between them.

Unlike Baulch et al. (2007) and van de Walle and Gunewardena (2001), we do not include commune-level fixed effects because the number of households in each commune is very small for VHLSS series (e.g., there are only three households with consumption data in most communes) and a relatively large number of communes are included in the data. Therefore, we choose to include the provincial-level fixed effects to address potential heterogeneity across locations. For the VLSS series, this has a similar effect to the commune-level fixed effects because there are only a few observed communes in each province. To allow for arbitrary correlation across households in a commune, we report standard errors clustered at the commune level.

As the comparisons of regression results in Tables 2 and 3 across years and between the ethnic minority and majority groups indicate, there are cases where the sign of the coefficient is different between the ethnic minority and majority groups or dependent on the survey round. For example, Table 2 shows that the coefficient on female-headed household is positive and significant in both 1993 and 2002 for the ethnic majority group, whereas it is insignificant for both years and negative in 2002 for the ethnic minority group. Table 3 also shows that the sign of the coefficient on female-headed household is dependent on the ethnicity and survey round. However, when both positive and negative signs are observed for a given covariate, one of the two signs is statistically insignificant in all cases except for the area for rice and its squared term.⁶ Therefore, the regression results presented in Tables 2 and 3 suggest that the direction of impact of each covariate is consistent across the ethnic groups and over time.

However, the coefficients are quantitatively different between the ethnic minority and majority groups. For each year and each panel data set, we conducted a Wald test of equality of coefficients for all the reported regressors between the ethnic minority and majority groups. We strongly reject the null hypothesis of equality of coefficients in all regressions. This indicates the presence of persistent structural differences between the ethnic minority and majority groups.

⁵Unreported regression results are available upon request.

⁶One possible reason why the logarithmic consumption per capita became a convex function of land area for rice for the ethnic majority in recent years but not for the ethnic minority group may reflect the increased capital intensity in the rice production for the ethnic majority group.

TABLE 2 Regression results by the ethnic minority/majority groups for the years 1993 and 2002

Year Minority/Majority	1993		2002	
	Minority	Majority	Minority	Majority
Female-headed household	0.018 (0.050)	0.069* (0.036)	-0.025 (0.083)	0.109** (0.046)
Age of household head	0.017* (0.010)	0.014*** (0.005)	0.020* (0.010)	-0.006 (0.008)
Age of household head sq.	-1.62e-4 (1.05e-4)	-1.10e-4** (5.59e-5)	-2.11e-4* (1.08e-4)	7.29e-5 (7.89e-5)
Head is married	0.109* (0.057)	0.120*** (0.037)	0.133* (0.078)	0.170*** (0.045)
Logarithmic household size	-0.428*** (0.050)	-0.258*** (0.043)	-0.453*** (0.056)	-0.222*** (0.043)
Ratio of children	-0.329*** (0.087)	-0.444*** (0.060)	-0.474*** (0.106)	-0.545*** (0.074)
Ratio of elderly	0.088 (0.166)	-0.122** (0.059)	0.077 (0.185)	-0.348*** (0.077)
Primary completed	0.057 (0.043)	0.086** (0.034)	0.066* (0.038)	0.082** (0.039)
Lower secondary completed	0.171*** (0.063)	0.199*** (0.039)	0.241*** (0.052)	0.182*** (0.040)
Upper secondary completed	0.156* (0.091)	0.263*** (0.045)	0.225*** (0.069)	0.281*** (0.050)
University or higher	0.494* (0.260)	0.351*** (0.074)	0.586** (0.232)	0.563*** (0.084)
Head employed in primary sec.	0.018 (0.074)	0.014 (0.045)	0.030 (0.084)	-0.039 (0.056)
Head employed in secondary sec.	0.176* (0.094)	0.121* (0.064)	0.185 (0.150)	0.066 (0.065)
Head employed in tertiary sec.	0.154 (0.164)	0.251*** (0.057)	0.269** (0.125)	0.185*** (0.067)
Area for rice	3.96e-5** (1.54e-5)	4.22e-5*** (8.63e-6)	1.44e-5** (6.61e-6)	-3.46e-5*** (1.05e-5)
Area for rice squared	-1.21e-9 (1.04e-9)	-9.27e-10*** (2.78e-10)	-5.65e-11 (1.37e-10)	2.64e-9*** (8.09e-10)
Area for other annual	6.92e-6 (1.14e-5)	1.12e-5 (9.53e-6)	5.39e-7 (5.11e-6)	1.56e-5* (8.01e-6)
Area for other annual sq.	3.05e-10 (4.42e-10)	-1.16e-10 (1.60e-10)	1.31e-10** (6.34e-11)	-6.20e-11 (1.63e-10)

(Continues)

TABLE 2 (Continued)

Year Minority/Majority	1993		2002	
	Minority	Majority	Minority	Majority
Area for annual industrial	2.11e-4*** (3.05e-5)	-1.82e-5 (2.39e-5)	-7.80e-7 (2.76e-5)	5.77e-6 (1.59e-5)
Area for annual industrial sq.	-2.76e-8*** (4.47e-9)	1.76e-9 (1.68e-9)	8.05e-10 (2.82e-9)	1.71e-10 (3.15e-10)
Area for perennial industrial	6.44e-5 (5.36e-5)	2.53e-6 (2.76e-5)	2.24e-5*** (5.45e-6)	6.69e-7 (7.35e-6)
Area for perennial industrial sq.	-2.06e-8 (1.65e-8)	8.35e-11 (1.20e-9)	-3.81e-11*** (9.43e-12)	6.00e-11 (1.25e-10)
Area for fruit trees	-0.001 (8.84e-4)	3.68e-5 (6.78e-5)	6.88e-5 (4.97e-5)	2.47e-5 (1.72e-5)
Area for fruit trees sq.	1.10e-7 (9.62e-8)	-2.98e-9 (7.76e-9)	-3.49e-9 (5.65e-9)	-4.02e-10 (5.29e-10)
Data	Panel VLSS 1993–1998		Panel VHLSS 2002–2004	
R ²	0.6631	0.3492	0.5279	0.3072
Number of observations	361	1,609	402	1,326

Notes: Provincial-level fixed effects are included in all regressions. Standard errors clustered at the community level are in parenthesis.

*, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

We have also carried out a Wald test of the equality in the change of coefficients between the initial and terminal time points between the two ethnic groups. Here, we find that the ethnic difference in the returns to characteristics changed during the 1993–1998 (significant at 1%), 2006–2008 (significant at 1%) and 2012–2014 (significant at 5%) periods. Therefore, while the structural difference between the two ethnic groups persisted, the nature of the structural difference changed significantly during these periods.

It should be noted here that Equation (1) makes an implicit assumption that the coefficients β may be different between the ethnic minority and majority groups but they are the same within each of these groups, because our main focus is on the gap between the ethnic minority and majority groups. Further, our decomposition requires assumptions about the way the coefficients change over time as Appendix B makes clear. Therefore, the assumption that the coefficients are the same within each ethnic group allows us to keep minimum the set of assumptions on the trajectory of the coefficients. Finally, under our assumptions, OLS regressions has an added advantage over other methods relying on local weighting that the former tends to produce more accurate estimates.

Nevertheless, because the structural gap within each group is also of interest, we carried out unconditional quantile regressions proposed by Firpo, Fortin, and Lemieux (2009), which allows us to identify the heterogeneity of the impact of changes in the distribution of the covariates across different quantiles. The results of these regressions show that the signs and sizes of the coefficients are generally comparable to those reported in Tables 2 and 3. Further discussion on these results is provided in Appendix S1. Below, we examine the returns to demographic characteristics, education, sector of employment and cultivated land areas separately.

TABLE 3 Regression results by the ethnic minority/majority groups for the years 2008 and 2014

Year Minority/Majority	2008		2014	
	Minority	Majority	Minority	Majority
Female-headed household	-0.037 (0.070)	0.075* (0.039)	0.028 (0.081)	-0.065 (0.057)
Age of household head	0.005 (0.010)	0.003 (0.007)	0.029*** (0.010)	-0.005 (0.009)
Age of household head sq.	-6.72e-5 (1.05e-4)	-3.32e-5 (6.96e-5)	-2.79e-4*** (1.03e-4)	4.16e-5 (7.72e-5)
Head is married	0.268*** (0.090)	0.125** (0.049)	0.092 (0.091)	0.047 (0.066)
Logarithmic household size	-0.435*** (0.057)	-0.166*** (0.038)	-0.403*** (0.070)	-0.207*** (0.043)
Ratio of children	-0.248** (0.109)	-0.552*** (0.076)	-0.373*** (0.129)	-0.420*** (0.088)
Ratio of elderly	0.070 (0.139)	-0.281*** (0.074)	-0.343* (0.177)	-0.299*** (0.079)
Primary completed	0.240*** (0.044)	0.088** (0.040)	0.125*** (0.044)	0.133** (0.052)
Lower secondary completed	0.242*** (0.051)	0.229*** (0.041)	0.210*** (0.057)	0.265*** (0.048)
Upper secondary completed	0.426*** (0.075)	0.342*** (0.044)	0.286** (0.114)	0.455*** (0.054)
University or higher	0.636*** (0.115)	0.431*** (0.072)	0.549* (0.319)	0.560*** (0.108)
Head employed in primary sec.	-0.039 (0.095)	-0.023 (0.049)	-0.102 (0.130)	0.050 (0.050)
Head employed in secondary sec.	-0.014 (0.150)	0.079 (0.057)	0.053 (0.144)	0.069 (0.052)
Head employed in tertiary sec.	0.092 (0.121)	0.119* (0.065)	0.197 (0.189)	0.205*** (0.057)
Area for rice	1.21e-5** (5.73e-6)	-1.85e-5*** (4.04e-6)	7.12e-6 (9.73e-6)	-1.10e-5 (7.30e-6)
Area for rice squared	-6.32e-11 (7.31e-11)	4.41e-10*** (9.43e-11)	-1.06e-11 (3.92e-10)	4.39e-10 (3.00e-10)
Area for other annual	3.76e-6 (4.66e-6)	-3.35e-6 (7.19e-6)	9.44e-7 (8.58e-6)	-4.80e-6 (1.72e-5)
Area for other annual sq.	3.76e-11 (4.75e-11)	7.59e-11 (1.17e-10)	-4.86e-11 (2.58e-10)	1.82e-10 (7.08e-10)

(Continues)

TABLE 3 (Continued)

Year Minority/Majority	2008		2014	
	Minority	Majority	Minority	Majority
Area for annual industrial	1.40e-5 (2.15e-5)	2.16e-5 (1.34e-5)	5.28e-5 (3.97e-5)	6.22e-6 (1.81e-5)
Area for annual industrial sq.	1.38e-9 (1.56e-9)	-7.23e-10 (5.30e-10)	-1.84e-9 (4.79e-9)	-5.28e-11 (4.86e-10)
Area for perennial industrial	4.23e-5*** (1.44e-5)	1.24e-5 (8.69e-6)	9.86e-5*** (2.17e-5)	7.93e-6 (1.08e-5)
Area for perennial industrial sq.	-1.21e-9*** (3.89e-10)	-3.57e-11 (2.89e-10)	-3.60e-9*** (1.16e-9)	-4.96e-11 (3.32e-10)
Area for fruit trees	4.69e-5 (5.81e-5)	5.24e-5* (3.05e-5)	1.48e-5 (1.20e-4)	-1.58e-5 (2.57e-5)
Area for fruit trees sq.	-6.23e-9 (5.80e-9)	-2.51e-9 (1.95e-9)	8.92e-9 (1.64e-8)	1.69e-9 (1.36e-9)
Data	Panel VHLSS 2012–2014		Panel VHLSS 2012–2014	
R^2	0.4720	0.3347	0.4584	0.2967
Number of observations	482	1,383	547	1,182

Notes: Provincial-level fixed effects are included in all regressions. Standard errors clustered at the community level are in parenthesis.

*, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

4.1 | Demographic characteristics

Tables 2 and 3 show that larger households tend to be poorer for both the minority and majority groups. The coefficient on logarithmic household size is negative and significant in all reported and unreported regressions at the 1% level. Furthermore, there is also a significant difference between the ethnic minority and majority groups in the impact of household size. The estimated coefficient for the ethnic minority groups is roughly -0.4 for all years. On the other hand, the corresponding figure for the ethnic majority group is around -0.2 . Therefore, the negative impact of household size on the logarithmic consumption for the minority households is twice as large as that for majority households.

Higher concentration of children (aged 14 and below) tends to lower the standard of living for both minority and majority households. This effect appears to be similar between minority and majority households. On the other hand, higher concentration of elderly people (aged 65 and above) has a significantly negative impact on the standard of living only for the majority households. For the minority households, it has no effect for most years and marginally negative effect only for the years 2012 (unreported) and 2014.

4.2 | Education

The coefficient estimates on the household head's education reported in Tables 2 and 3 underscore the importance of human capital investment. All coefficients are positive, most coefficients are

statistically and economically significant, and higher levels of education tend to have a larger coefficient. In particular, the coefficient on the completion of lower secondary education is always significant in reported and unreported regressions and mostly significant at the 1% level.

It is not clear from these tables whether the returns to education are the same between the ethnic minority and majority groups. Therefore, we report in Table 4 the difference in the returns to education and their changes over a given time period between the majority and minority groups. In each segment of Table 4 sandwiched by two solid lines below the first row, we report the differences in regression coefficients between the ethnic majority and minority groups at the initial period t_0 and terminal period t_1 or $\Delta_{t_i} \equiv \beta_{t_i}^1 - \beta_{t_i}^0$ for $i \in \{0, 1\}$. Below the dashed line, we report the DiD, which is $\Delta_{t_1} - \Delta_{t_0}$ using the notations introduced above. For example, Table 4 shows that the returns to lower secondary education of the household head for the ethnic majority group are higher than those for the ethnic minority group by 0.028 (= 0.199–0.171).

While most of the differences are statistically insignificant in Table 4, two observations are worth highlighting. First, the only negative and significant differences are the difference in 2008 for primary education and DiD between 2006 and 2008 for primary and lower secondary education. The difference in returns to primary education became less negative in 2010 and positive in 2012 and 2014. Thus, the returns to primary education for the ethnic minority relative to that for the ethnic majority group were highest around 2008.

TABLE 4 Difference in returns to education between the ethnic majority and minority groups by different education level

Reference	Primary	Lower sec	Upper sec	Univ. or higher
1993	0.029	0.028	0.107	–0.142
1998	0.028	–0.042	0.010	–0.061
DiD	0.000	–0.070	–0.097	0.081
2002	0.016	–0.058	0.055	–0.024
2004	0.006	–0.037	–0.073	–0.058
DiD	–0.010	0.022	–0.129	–0.035
2004	–0.082	–0.064	0.023	–0.030
2006	–0.016	0.109*	0.005	0.090
DiD	0.066	0.173**	–0.018	0.120
2006	–0.017	0.137**	0.023	–0.092
2008	–0.152***	–0.013	–0.084	–0.205
DiD	–0.135*	–0.150*	–0.107	–0.113
2010	–0.053	0.014	–0.014	0.367**
2012	0.089	0.004	0.039	0.199
DiD	0.142*	–0.010	0.053	–0.168
2012	0.091	0.141**	0.149	0.184
2014	0.008	0.055	0.169	0.010
DiD	–0.083	–0.086	0.020	–0.174

Notes: Positive [negative] difference implies that the returns to the employment sectors for the ethnic majority group are higher [lower] those that for the ethnic minority group.

*, ** and *** denote statistical significance at 10%, 5% and 1% levels, respectively.

Second, DiD is negative between 2006 and 2008 at all levels of education and the point estimates were highly negative in 2008. Therefore, the difference in the returns to education clearly shifted in favour of the ethnic minority group during this time period. However, in all the other time periods, the sign of DiD is mixed and thus it is not clear whether the changes in returns to education overall contributed to the increase or decrease of the development gap between the ethnic minority and majority groups.

While we are unable to identify the sources of changes in returns to education, several factors are likely to be at play. First, the supply effect is likely to be relevant. As the supply of skilled workers increases, the premium paid for an educated labour force tends to drop. The decline in the difference in returns to primary education may be partly explained by this effect. As Table 1 shows, the proportion of the majority household heads with at least primary education has increased from 78.3 (= 100–21.7)% to 89.3 (= 100–10.7)% between 1993 and 2014 in our sample. On the other hand, the corresponding proportion of the minority household heads did not change much. Therefore, if the labour markets for the ethnic minority and majority groups are segregated, the decline in the scarcity of majority household heads who have completed primary education may have contributed to the decline in the difference in the returns to primary education.

The supply effect may also work through the selection of people who get educated. To see this point, suppose that those who have the highest returns to education (because they have the highest innate ability, for example) are the ones who receive education first. Then, as education spreads to a higher share of each cohort, the average returns to education drop.

The second relevant factor is the quality of education. The government has stepped up efforts to include the ethnic minority group in the education system and improve the quality of education for the ethnic minority households. If the gap in the quality of education between the ethnic minority and majority groups has narrowed, it may show up as a negative change in the difference in the returns to education.

Finally, technological changes and international trade are also likely to have played a role. It is plausible that the ethnic majority group was better positioned to take advantage of new opportunities created by international trade and new technologies than the ethnic minority group. If these opportunities eventually spill over to the ethnic minority group, the positive differences in returns to education would revert to zero eventually. While we cannot draw strong conclusions here, this explanation appears to have some relevance because the differences tend to get smaller after a relatively large difference is observed.

4.3 | Employment sector

Tables 2 and 3 show that the sector of employment matters for the standards of living. The coefficients on the tertiary sector tend to be the highest followed by the secondary and primary sectors in this order. This is true for both the ethnic minority and majority groups. As Table 1 shows, the proportions of the majority household heads employed in the secondary and tertiary sectors exceed those of the minority household heads. Therefore, the difference in the composition of sectors of employment between the ethnic minority and majority groups also explains the gap in the average logarithmic consumption per capita between the ethnic minority and majority groups.

As with the returns to education, we compute the differences between the ethnic minority and majority groups in the returns to the sector of employment and their change over time. Table 5 reports these differences. A notable observation from Table 5 is that DiD for the period between 2012 and 2014 is positive for all sectors and significant for the primary and tertiary sectors. This

TABLE 5 Difference in the returns to primary, secondary and tertiary sectors of employment for household head between the ethnic majority and minority groups

Reference	Primary	Secondary	Tertiary
1993	−0.004	−0.055	0.097
1998	−0.086	0.067	−0.064
DiD	−0.081	0.121	−0.161
2002	−0.070	−0.119	−0.084
2004	0.049	−0.137	−0.035
DiD	0.119	−0.018	0.049
2004	0.328***	0.159	0.233*
2006	0.246**	0.182	0.180
DiD	−0.082	0.023	−0.053
2006	0.129	−0.005	0.143
2008	0.016	0.093	0.027
DiD	−0.113	0.098	−0.116
2010	−0.048	−0.105	−0.354***
2012	−0.159	−0.187	−0.416**
DiD	−0.111	−0.082	−0.061
2012	−0.183	−0.059	−0.494***
2014	0.152	0.016	0.008
DiD	0.335**	0.075	0.503**

Notes: Positive [negative] difference implies that the returns to characteristic for the ethnic majority group are higher [lower] than that for the ethnic majority group.

*, ** and *** denote statistical significance at 10%, 5% and 1% levels, respectively.

is particularly concerning because the difference in the returns to employment is contributing to the gap between the ethnic minority and majority groups in recent years.

4.4 | Cultivated land areas

In Tables 2 and 3, the coefficient on the land area cultivated for rice tends to be statistically significant. We also find that the land area cultivated for perennial industrial crops tends to be an important determinant for the ethnic minority group, but this is not the case for the ethnic majority group.

Because we include second-order terms, the difference in the returns to land areas and their change over time between the ethnic groups vary with the amount of land that households have. Therefore, for each category of land and for each panel data set, we evaluate the marginal effect of land area at the median among those who have a strictly positive amount of area in that category. In other words, we are evaluating the marginal effect of each category of land for those who have a typical amount of land in that category.

Table 6 reports the first difference (i.e., difference between minority and majority group) in the marginal effect and the DiD. For example, the marginal effect for the ethnic majority groups of land cultivated for rice at the median land size among those who have any such land was higher

TABLE 6 Difference in the marginal effects of land on the logarithmic consumption per capita between the ethnic majority and minority groups evaluated at the median land area among those who have some land in a given category

Period	Rice	Other annual	Annual industrial	Perennial industrial	Fruit trees
1993	5.03e-6	3.44e-6	-2.03e-4***	-4.70e-5	1.05e-3
1998	-2.17e-5**	-1.12e-5	-6.10e-5***	-2.26e-5	1.20e-4
DiD	-2.67e-5**	-1.46e-5	1.42e-4***	2.44e-5	-9.29e-4
2002	-2.81e-5***	1.47e-5*	5.92e-6	-2.13e-5**	-4.28e-5
2004	-3.30e-5***	-7.08e-6	1.73e-5	-1.42e-5	-1.33e-4***
DiD	-4.91e-6	-2.18e-5	1.14e-5	7.19e-6	-8.99e-5
2004	-3.41e-5***	1.56e-6	-2.01e-5	-4.65e-6	-4.10e-5
2006	-1.29e-5***	-1.42e-5	7.17e-6	-2.75e-5	8.65e-5
DiD	2.12e-5***	-1.58e-5	2.73e-5	-2.29e-5	1.28e-4*
2006	-2.61e-5***	-3.86e-6	-1.55e-5	-2.32e-5	2.95e-5
2008	-2.35e-5***	-7.02e-6	5.35e-6	-2.41e-5*	6.84e-6
DiD	2.62e-6	-3.16e-6	2.09e-5	-9.26e-7	-2.27e-5
2010	-2.31e-5***	6.84e-6	-7.24e-5**	1.14e-6	-7.94e-5
2012	-2.18e-5**	-7.33e-6	-6.63e-6	8.08e-6	-3.35e-5
DiD	1.38e-6	-1.42e-5	6.58e-5*	6.94e-6	4.59e-5
2012	-1.46e-5	-2.21e-5	1.59e-5	-1.21e-5	6.34e-5
2014	-1.52e-5*	-5.29e-6	-4.45e-5	-6.93e-5***	-3.35e-5
DiD	-6.59e-7	1.68e-5	-6.04e-5	-5.72e-5***	-9.70e-5

Notes: Positive [negative] difference implies that the returns to marginal land for the ethnic majority group is higher [lower] than that for the ethnic majority group.

*, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

by 5.03×10^{-6} than that for the ethnic minority group in 1993. By 1998, the difference dropped by 2.67×10^{-5} to -2.17×10^{-5} .

Table 6 also shows that the marginal impact of land cultivated for rice is significantly negative for most years. Therefore, a marginal land plot for rice given to a household that is already cultivating a land plot of typical size for rice cultivation would lead to a higher increase in the standards of living if the household head is ethnic minority than it would if the head is ethnic majority.

5 | POVERTY DECOMPOSITION ANALYSIS

The preceding regression analysis shows the presence of persistent but changing structural differences between the ethnic minority and majority groups. However, it does not tell clearly how this structural difference translates into the development gap between the ethnic groups. A popular approach to this issue is the Oaxaca–Blinder decomposition. In Vietnam, the Oaxaca–Blinder decomposition and its variant were applied to household consumption by Baulch et al. (2007), Nguyen et al. (2007), and van de Walle and Gunewardena (2001) and to wage by Phan and Coxhead (2013).

While it is straightforward to apply the Oaxaca–Blinder decomposition to our data, we choose to use a variant of a regression-based decomposition proposed by Fujii (2015) using a panel data set, which is summarised in Appendix B.⁷ We apply it to the poverty gap measure, or the average shortfall from the poverty line as a share of the poverty line, where the shortfall is counted as zero if one is already above the poverty line. We report the results under \$1.90-a-day poverty line, but the results discussed in this paper are qualitatively similar even when the poverty severity (a.k.a. squared poverty gap) or Watts poverty measures are used or when \$3.10-a-day poverty line is used.

The method proposed by Fujii (2015) has at least three advantages over the popular Oaxaca–Blinder decomposition. First, the Oaxaca–Blinder decomposition does not work well in Vietnam when the location-specific fixed-effects terms are included. This is true whether the fixed effects are defined at the commune level or provincial level because some commune and provinces have only minority or majority households (but not both). This situation is not specific to Vietnam because it is common to see spatial segregation of different ethnic groups in other countries. On the other hand, once we aggregate the contributions from location-specific fixed effects, we are able to apply the decomposition by Fujii (2015) to the cases where some locations have only minority or majority households.

Second, because our focus is on development gap, we are most interested in the lower tail of consumption distribution. This point is important because a sizable fraction of households may be experiencing declining consumption, even when the average logarithmic consumption is increasing. Consistent with this possibility, United Nations Development Programme (2013) reports that the average income of households increased between 2007 and 2012 but 46% of households experienced a decrease in real income. Therefore, the analysis of the mean logarithmic consumption difference between the ethnic minority and majority groups, which is essentially what the Oaxaca–Blinder decomposition achieves, may not be helpful for the understanding of the important contributing factors to persistently high poverty for the ethnic minority group relative to the ethnic majority group. While it is also possible to analyse conditional quantiles instead of the mean using quantile regressions (Machado & Mata, 2005), the Oaxaca–Blinder decomposition does not tell us how much the changes in the structural parameters and distributions of covariates contributed to the observed change in poverty.

Finally, the Oaxaca–Blinder decomposition requires us to specify the reference group, whose choice is not obvious and often arbitrary. Furthermore, this choice can affect the decomposition results substantially. In contrast, Fujii (2015) does away with this choice, because we first decompose the change in poverty for ethnic minority and majority groups separately and then take a difference between the two groups for each factor.

Table 7 reports the decomposition of the poverty gap measure. At the top part of each column, we report the time period of decomposition, the difference in poverty gap at the initial and terminal years between the ethnic minority and majority groups, and the change in the difference in poverty gap. For example, Table 7 shows that the poverty gap for the ethnic minority was higher than that for the ethnic majority by 0.0843 in 1993. This difference increased by 0.0575 to 0.1418 between 1993 and 1998. Consistent with the discussion in the previous section, the absolute difference in poverty gap between the ethnic minority and majority groups widened between 1993 and 2014.

Below these rows, we report the contribution of various factors to the change in the difference in poverty gap between the ethnic groups. The *X* subcolumn represents the contribution due to the differences in the change in the distribution of covariates between the two groups, which we call

⁷Wan (2002) proposed a regression-based Shapley decomposition of inequality measures (See also Wan (2004); Wan and Zhou (2005) for empirical applications), which could be modified for poverty decomposition. However, Shapley decomposition only considers sequential changes of underlying factors. Fujii (2015), on the other hand, allows all the factors to change simultaneously. Because sequential change is unlikely to hold in practice, we choose to follow Fujii (2015).

TABLE 7 Decomposition of the difference in the poverty gap measure between the ethnic majority and minority groups

Time period	1993–1998		2002–2004		2004–2006		2006–2008		2010–2012		2012–2014	
	X	S	X	S	X	S	X	S	X	S	X	S
Initial poverty difference	-0.0843		-0.1399		-0.1841		-0.1875		-0.1771		-0.1478	
Terminal poverty difference	-0.1418		-0.1779		-0.1802		-0.1819		-0.1360		-0.1146	
Change in difference	-0.0575		-0.0380		0.0039		0.0056		0.0411		0.0331	
	X	S	X	S	X	S	X	S	X	S	X	S
Demographic	-0.0145	-0.0402	-0.0109	-0.0766	-0.0046	-0.2354	-0.0042	0.0774	0.0091	0.0052	0.0010	0.0814
Education	0.0004	0.0093	0.0015	-0.0022	0.0001	-0.0342	0.0063	0.0392	0.0014	-0.0063	0.0000	0.0059
Employment	-0.0030	0.0402	0.0045	-0.0457	-0.0009	0.0338	0.0034	0.0499	-0.0033	0.0384	0.0006	-0.0695
Land	-0.0074	0.0518	-0.0020	0.0384	0.0306	-0.0798	-0.0006	-0.0320	-0.0062	0.0041	-0.0010	-0.0026
Provincial FE	0.0000	-0.0954	0.0000	0.0489	0.0000	0.2921	0.0000	-0.1254	0.0000	0.0054	0.0000	0.0193
Subtotal	-0.0246	-0.0343	-0.0070	-0.0373	0.0252	-0.0235	0.0048	0.0091	0.0009	0.0467	0.0005	0.0345
Residual	0.0013		0.0004		0.0059		-0.0076		-0.0046		0.0004	
Weight	0.0001		0.0059		-0.0037		-0.0007		-0.0019		-0.0023	

Note: The difference is calculated as the majority figure minus minority figure. See Appendix B for relevant details of the implementation of the decomposition by Fujii (2015).

“covariate component” for simplicity. The S subcolumn represents the contribution due to the differences in the change in return to characteristics between the two groups, which we call “structural component.” At the bottom of each column, the contributions of the changes in the residual term (R in Appendix B) and the sample weights (W) are reported. These components turn out to be small in our study.

Table 7 shows that the absolute difference in poverty gap widened until 2004, but it has generally narrowed afterwards. Except for the period between 2004 and 2006, both the subtotal of covariate (X) and structural (S) components have the same sign and the absolute value of the subtotal for the structural component exceeds that for the covariate component. Therefore, a significant portion of the changes in the difference between the ethnic minority and majority groups in their poverty gap measures are driven by the structural change over time. This finding echoes previous static Oaxaca–Blinder decomposition studies by Baulch et al. (2007) and van de Walle and Gunewardena (2001), which also underscore the importance of the structural component.

In addition to decomposing into covariate and structural components, we also report the breakdown of these components into different factors. Because the location of a household does not change (see Appendix A), the provincial fixed effect component is equal to zero for all time periods.⁸ Similarly, because very few household heads change the educational attainment, the covariate component is close to zero.

It should be pointed out here that the structural components reported in Table 7 are consistent with the analysis of returns to characteristics discussed above. For example, we saw that the difference in returns to education shifted in favour of the ethnic minority group between 2006 and 2008. Consistent with this, the structural component for education is positive in this time period, indicating that the difference of the poverty gap measure for the ethnic minority group from that of the ethnic majority group has narrowed due to the structural change in returns to education.

When we take the combined effects of both covariate and structural components, demographic characteristics, sector of employment and provincial fixed effects are generally among important factors. On the other hand, education is not particularly important in explaining the difference in the changes in poverty between ethnic minority and majority groups. Land appears to have been important until around 2006. However, in more recent years, the importance of land seems to have diminished.

6 | ETHNICITY AS A SOURCE OF INEQUALITY OF OPPORTUNITY

The results presented in the previous section focus on those who are under the poverty line. It is, however, of interest to see if our empirical results qualitatively change when the upper tails are also taken into account. In this section, we underscore how our poverty decomposition analysis relates to the inequality decomposition analysis used in the inequality of opportunity literature.

As mentioned in Section 1, ethnicity can be viewed as a circumstance variable, because the ethnic group individuals belong to is something that they should not be held responsible for. Hence, it is possible to apply the method in the literature on inequality of opportunity to analyse the development gap. Because our focus is on ethnicity, we do not consider other circumstance variables (e.g., sex) that may lead to inequality of opportunity.

⁸While the difference in the migration patterns between the urban and rural areas in general affects the development gap between the ethnic minority and majority groups, it is not a concern for us because the same set of non-migratory households is used for our decomposition analysis.

TABLE 8 Overall and between-group Theil-L inequality measures

Initial inequality						
Initial year	1993	2002	2004	2006	2010	2012
Overall	0.0881	0.1139	0.1249	0.1282	0.1846	0.1605
Between group	0.0033	0.0077	0.0124	0.0138	0.0366	0.0332
Terminal inequality						
Terminal year	1998	2004	2006	2008	2012	2014
Overall	0.1008	0.1150	0.1278	0.1237	0.1610	0.1759
Between group	0.0075	0.0097	0.0124	0.0169	0.0313	0.0317
Change in inequality between the initial and terminal years						
Overall	0.0127	0.0011	0.0028	-0.0045	-0.0236	0.0154
Between group	0.0042	0.0020	0.0001	0.0030	-0.0053	-0.0015

Following Checchi and Peragine (2010) and Ferreira and Gignoux (2011), we decompose the Theil-L index (a.k.a. mean log deviation). Because we have the ethnic minority and majority groups as subgroups and sufficient observations in each group, we only consider the non-parametric method. Then, the absolute level of inequality of opportunity due to the ethnicity can be obtained as the between-group component in the standard decomposition of the Theil-L index by population subgroups. In our context, this between-group inequality can be interpreted as the reduced-form effect of ethnicity on inequality, which include the direct effect of the ethnicity as a circumstance variable and the indirect effect through efforts because circumstance variables may alter the efforts of individuals.

Table 8 provides the results of the inequality decomposition using the six panel data sets used in the earlier two sections. The top part of the table shows the overall inequality and between-group inequality as measured by the Theil-L index in the initial year, whereas the mid part of the table shows those in the terminal year. Therefore, Table 8 shows that the Theil-L index generally increased over time from 0.0881 in 1993 to 0.1759 in 2014. However, the importance of between-group inequality declined after the year 2010. This result is broadly consistent with the poverty decomposition results presented in Section 5, because the difference in poverty gap between the ethnic minority and majority groups also narrowed substantially after the year 2010.

7 | IS AID REACHING THE ETHNIC MINORITY GROUP?

The preceding analysis shows that there remains a sizable development gap between the ethnic minority and majority groups. Even though both the structural and covariate components have contributed favourably to narrow the develop gap after 2006, narrowing the development gap between the ethnic groups remains an important challenge in Vietnam.

To reduce poverty for the minority group and tackle the development gap between ethnic minority and majority groups, a number of state-level programmes have been implemented. For example, Programmes 132/134 are designed to reallocate land to those minority households that lacked land and Programme 135 (P135) targets poor and remote areas with a high concentration of ethnic minority households. Other interventions include cash transfer, health and education programmes.

The mere existence of such programmes does not guarantee that the poor minority households will benefit from them. Further, even if the resources for a specific programme are reaching minority households, the total amount of resources given to minority households may not exceed that

TABLE 9 Difference between the actual and predicted proportions of aid recipients for health, education and credit programmes

Year	Health Prog.	Educ. Prog.	Credit Prog.
2002	0.084		
2004	0.356	0.468	0.254
2006	0.301	0.250	0.154
2008	0.026	0.160	0.339
2010	0.566	0.238	0.396
2012	0.735	0.382	0.366
2014	0.726	0.222	0.178

Notes: Health assistance includes support for the purchase of health insurance card and fee remission from medical examination and treatment for the poor. Education assistance includes exemption from tuition fees for the poor and policy-based scholarships. Credit assistance includes preferential credit for the poor.

for majority households. For example, Nguyen, Phung, Phung, Ngoc, and Westbrook (2012) find that the treatment communes under the second phase of P135 received significantly more P135 funding than the comparison communes but that the total amount of resources given to P135 commune is on average no higher than that for the rest of the communes. This is because the provincial government reduces resources given to P135 communes.

Therefore, we check whether aid resources given to the minority household exceeds what one would expect had the majority household had a similar level of poverty using all rounds of VHLSS data sets (VLSS series do not contain relevant information). To be specific, we compute the following quantity R for each survey round:

$$R \equiv q^0 - [r^0 q_1^1 + (1 - r^0) q_0^1],$$

where q^0 is the proportion of aid recipients within the ethnic minority group, r^0 is the poverty rate among the minority group, and q_0^1 and q_1^1 are, respectively, the proportions of aid recipients among nonpoor and poor in the ethnic majority group in the survey year. Therefore, R can be interpreted as the difference between the actual probability of receiving assistance (q^0) and its predicted probability based on the poverty rate of the ethnic minority and probability of receiving assistance given the poverty status (expression within the square bracket).

Table 9 reports the value of R calculated with a \$1.90-a-day poverty line for the three most common types of assistance that households receive. As the table shows, the value of R is all positive. This point remains true even when we include other forms of assistance such as land, housing and vocational training, or when we use a \$3.10-a-day poverty line. Therefore, we have no evidence that the ethnic majority households are receiving more resources than ethnic minority households given the households' poverty status.

8 | DISCUSSION

Development gap between the ethnic minority and majority groups is a commonly observed issue around the globe. It is observed not only in Vietnam but also in many other multiethnic developing countries such as Bolivia, Central African Republic, China, Ecuador, Gabon, Guatemala, India, Laos, Mexico and Peru. In developed countries such as Australia, New Zealand and the US, the fact that a sizable gap in the living standards remains between indigenous and white people is well known.



Vietnam is arguably among the most interesting countries to study. With the successful pro-poor economic growth and impressive poverty reduction over more than two decades, one may expect that the development gap between the ethnic minority and majority groups has narrowed in Vietnam. However, our analysis of a series of household surveys indicates that this is unfortunately not the case. The gap persisted and even widened for some development indicators between 1993 and 2014.

Our regression analysis indicates that there is a persistent structural difference between the ethnic minority and majority groups over our study period. We find that the returns to the land area for rice production for the ethnic minority group were significantly lower than that for the ethnic majority group for most years, which may reflect, among others, poorer quality of land and lower technology for the ethnic minority group. Other than this, we did not find persistent difference in returns to education, employment sector and land between the ethnic minority and majority groups. Therefore, the persistent structural difference cannot be attributed to persistent difference in returns to one or a few characteristics. Rather, the way in which the structural difference matters for the development gap changed over time.

To see whether the differences in the distribution of covariates is a more important driver of the development gap than the differences in the returns to characteristics, we also apply a new decomposition technique proposed by Fujii (2015) to the difference in poverty gap between the ethnic minority and majority groups. This method allows us to focus on a lower tail of distribution unlike the popular Oaxaca–Blinder decomposition. This feature is particularly important for analysing the persistent difference in poverty between the ethnic minority and majority groups.

Our decomposition analysis shows that the structural component is at least as important as the covariate component. We also find that the structural changes have contributed to the increase in the difference in poverty between the ethnic minority and majority groups until 2006, but the sign of its contribution reversed after 2006. Further, we find that demographic characteristics, sector of employment and the provincial fixed effects are among the most important factors that explain the changes in the poverty difference between the ethnic minority and majority groups.

The government of Vietnam has been well aware of the development gap across different ethnic groups, and a series of programmes have been implemented to address this issue. One important question therefore is whether the minority groups are indeed benefiting from these programmes. While there are some issues in the resource allocation (e.g., countervailing reduction in the resources from the provincial government for the targeted communes in national programmes), we did not find any evidence that the ethnic majority groups are disproportionately benefiting from the aid programmes.

This, of course, does not mean that further improvement in targeting is unnecessary or impossible. However, what is critically needed is to rethink the way assistance is delivered. The fact that the decades of assistance did not lead to a visible narrowing of development gap between the ethnic minority and majority groups underscores this necessity. As some studies reviewed in Section 2 pointed out, there is a mismatch between what the central government thinks is needed and what the intended recipients really need. By addressing this mismatch and making policy more minority-appropriate, the development gap could be narrowed without spending significantly more resources for the aid.

While critical reassessment of the current policies and programmes is necessary, we do not need to be pessimistic as we can already see a silver lining. There are indications that paternalistic decision making on minority policies by the central government gradually changed towards a more open and inclusive one over time.⁹ Our finding that the provincial effects worked favourably

⁹For example, a decree to strengthen the Committee for Ethnic Minority Affairs (Decree 53/2004/ND-CP) was adopted in 2004. In 2013, the Ethnic Minority Affairs Strategy Through 2020 (Decision 449/QĐ-TTg), which addresses some of the mismatch issues mentioned above, was approved.

towards the ethnic minorities since 2006 may reflect this change. With further improvement in the minority policy and overall economic growth in Vietnam, we hope that more minority households will be lifted out of poverty and the gap between the ethnic minority and majority groups will be narrowed.

ACKNOWLEDGEMENTS

I gratefully acknowledge the receipt of SMU Research Grant (C244/MSS13E004) funded under the Singapore Ministry of Education's Academic Research Fund Tier 1 Programme for an earlier stage of this research. Yang Tao provided research assistance. I benefited from inputs from Ha T Doan, Charles Yuji Horioka, Jie Li, Yoko Niimi, Kang Park, Eric D. Ramstetter, Paul Saumik, Rebecca Valenzuela, Guanghua Wan, Eden Yu and two anonymous referees. An earlier version of this paper was presented at the ADBI-World Economy Workshop on Sources of Income Inequality in Asia held at the Asian Development Bank Institute.

ORCID

Tomoki Fujii  <http://orcid.org/0000-0002-9879-4835>

REFERENCES

- Aaberge, R., Mogstad, M., & Peragine, V. (2011). Measuring long-term inequality of opportunity. *Journal of Public Economics*, 95(3–4), 193–204.
- Baulch, B., Chuyen, T. T. K., Haughton, D., & Haughton, J. (2007). Ethnic minority development in Vietnam. *Journal of Development Studies*, 43(7), 1151–1176.
- Baulch, B., & Masset, E. (2003). Do monetary and nonmonetary indicators tell the same story about chronic poverty?: A study of Vietnam in the 1990s. *World Development*, 31(3), 441–453.
- Baulch, B., Nguyen, H. T. M., Phuong, P. T. T., & Pham, H. T. (2010). *Ethnic minority poverty in Vietnam* (Working Paper No. 169). Manchester, UK: Chronic Poverty Research Centre. Retrieved from <https://assets.publishing.service.gov.uk/media/57a08b0040f0b649740008e0/WP169.pdf>
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8, 436–455.
- Bourguignon, F., Ferreira, F. H. G., & Menéndez, M. (2007). Inequality of opportunity in Brazil. *Review of Income and Wealth*, 53(4), 585–618.
- Cecchi, D., & Peragine, V. (2010). Inequality of opportunity in Italy. *Journal of Economic Inequality*, 8, 429–450.
- Dang, H.-A. (2014). Vietnam: A widening poverty gap for ethnic minorities. In G. H. Hall, & H. Anthony (Eds.), *Patrimony indigenous peoples, poverty and development* (pp. 309–343). Cambridge: Cambridge University Press.
- Deaton, A., & Kozel, V. (2005). Data and dogma: The great Indian poverty debate. *World Bank Research Observer*, 20(2), 177–199.
- Economist* (2015). Out of sight. *Economist*, 415(8932), 36–37.
- Epprecht, M., Müller, D., & Minot, N. (2011). How remote are Vietnam's ethnic minorities?: An analysis of spatial patterns of poverty and inequality. *Annals of Regional Science*, 46, 349–368.
- Ferreira, F. H., & Gignoux, J. (2011). The measurement of inequality of opportunity: Theory and an application to Latin America. *Review of Income and Wealth*, 57, 622–657.
- Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3), 953–973.
- Fujii, T. (2015). *Poverty decomposition by regression: An application to Tanzania* (WIDER Working Paper 2015/102). Retrieved from World Institute for Development Economics Research website: <https://www.wider.unu.edu/sites/default/files/wp2015-102-.pdf>

- Glewwe, P., & Dang, H.-A. H. (2011). Was Vietnam's economic growth in the 1990s pro-poor? An analysis of panel data from Vietnam. *Economic Development and Cultural Change*, 59(3), 583–608.
- Hassine, N. B. (2011). Inequality of opportunity in Egypt. *World Bank Economic Review*, 26(2), 265–295.
- Haughton, D., & Haughton, J. (1997). Explaining child nutrition in Vietnam. *Economic Development and Cultural Change*, 45(3), 541–556.
- Lanjouw, J. O., & Lanjouw, P. (2001). How to compare apples and oranges: Poverty measurement based on different definitions of consumption. *Review of Income and Wealth*, 47(1), 25–42.
- Litchfield, J., & Justino, P. (2004). Welfare in Vietnam during the 1990s. *Journal of the Asia Pacific Economy*, 9(2), 145–169.
- Machado, J., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression models. *Journal of Applied Econometrics*, 20, 445–465.
- Marrero, G. A., & Rodríguez, J. G. (2013). Inequality of opportunity and growth. *Journal of Development Economics*, 104, 107–122.
- Marrero, G. A., Rodríguez, J. G., & van der Weide, R. (2016). *Unequal opportunity, unequal growth* (World Bank Policy Research Working Paper 7853). World Bank. Retrieved from World Bank website: <http://documents.worldbank.org/curated/en/473031476192890227/pdf/WPS7853.pdf>
- Nguyen, B. T., Albrecht, J. W., Vroman, S. B., & Westbrook, M. D. (2007). A quantile regression decomposition of urban-rural inequality in Vietnam. *Journal of Development Economics*, 83, 466–490.
- Nguyen, C., Phung, T., Phung, T., Ngoc, V., & Westbrook, D. (2012). *The impact of a national poverty reduction program on ethnic minorities in Vietnam: The lens of baseline and endline surveys* (MPRA Working Paper 50477). Retrieved from Mekong Development Research Institute website: https://mpra.ub.uni-muenchen.de/50477/1/MPRA_paper_50477.pdf
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14, 693–709.
- Phan, D., & Coxhead, I. (2013). Long-run costs of piecemeal reform: Wage inequality and returns to education in Vietnam. *Journal of Comparative Economics*, 41, 1106–1122.
- Pignataro, G. (2012). Equality of opportunity: Policy and measurement paradigms. *Journal of Economic Surveys*, 26(5), 800–834.
- Ramos, A. T. (2003). Vietnam. In C. Mackerras (Ed.), *Ethnicity in Asia* (pp. 157–173). Abingdon: RoutledgeCurzon.
- Ramos, X., & Van de Gaer, D. (2016). Approaches to inequality of opportunity: Principles, measures and evidence. *Journal of Economic Surveys*, 30(5), 855–883.
- Roemer, J. E. (1998). *Equality of opportunity*. Cambridge, MA: Harvard University Press.
- Tugault-Laeur, C., & Turner, S. (2011). Of rice and spice: Hmong livelihoods of diversification in the northern Vietnam Uplands. In J. Michaud, & T. Forsyth (Eds.), *Moving mountains: Ethnicity and livelihoods in Highland China, Vietnam, and Laos* (pp. 100–122). Vancouver, BC, Canada: UBC Press.
- United Nations Development Programme (2013). *Poverty situation analysis of ethnic minorities in Vietnam 2007–2012: Key findings from quantitative study*. Report prepared by Mekong Development Research Institute, United Nations Development Programme, Hanoi, Vietnam, December.
- van de Walle, D., & Gunewardena, D. (2001). Sources of ethnic inequality in Viet Nam. *Journal of Development Economics*, 65, 177–207.
- Vietnam News (2006). *Politics: Central highlanders opens 1st ethnic unity congress*. February 17.
- Wan, G. (2002). *Regression-based inequality decomposition: Pitfalls and a solution procedure* (WIDER Discussion Paper No. 2002/101). Retrieved from World Institute for Development Economics Research website: <https://www.wider.unu.edu/sites/default/files/dp2002-101.pdf>
- Wan, G. (2004). Accounting for income inequality in rural China: A regression-based approach. *Journal of Comparative Economics*, 32, 348–363.
- Wan, G., & Zhou, Z. (2005). Income inequality in rural China: Regression-based decomposition using household data. *Review of Development Economics*, 9(1), 107–120.
- World Bank (2009). *Ethnicity and development in Vietnam: Summary report*. Country Social Analysis, Social Development Unit, Sustainable Development Department, East Asia and Pacific Region, the World Bank.
- Writenet (2006). *Vietnam: Situation of indigenous minority groups in the Central Highlands*. A Writenet Report commissioned by United Nations High Commissioner for Refugees, Status Determination and Protection Information Section.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Fujii T. Has the development gap between the ethnic minority and majority groups narrowed in Vietnam?: Evidence from household surveys. *World Econ.* 2017;00:1–35. <https://doi.org/10.1111/twec.12578>

APPENDIX A

DATA APPENDIX

This study uses VLSS 1993 and 1998 as well as VHLSS 2002, 2004, 2006, 2008, 2010, 2012, and 2014, all of which were implemented by and can be obtained from the General Statistical Office of Vietnam. These are multitopic household surveys and cover a wide range of topics including household's demographic characteristics, education, health and employment of the household members, household assets, incomes, and consumption, and housing conditions. The exact set of questions included in the survey varies from round to round and some special modules are added to some rounds. We use those questions which were asked repeatedly across different rounds. We only use rural households in the loosely northern regions of Red River Delta, North East, North West, North Central, and Central Highlands regions in the eight-region classification.

To make the results comparable across years, we adopted a common definition of provinces, even though the definition of provinces changed slightly over time. To this end, we use the finest classification that could be identified from the data. According to this definition, Ha Tay Province in earlier definitions of provinces is included in Hanoi Province in our definition because Ha Tay was absorbed into Hanoi in 2008. Similarly, Dak Lak and Dak Nong provinces are taken as one province because Dak Nong was a part of Dak Lak province before 2004. For a similar reason, Lai Chau and Dien Bien provinces are taken as one province. Based on our definition of provinces, our sample includes 34 provinces out of a total of 61 provinces in Vietnam.

Both VLSS and VHLSS data sets contain panel households. Therefore, some of the households recorded in VLSS 1993 also appear in VLSS 1998. Similarly, VHLSS for years 2002–2008 and 2010–2014 have a partial rotating panel structure. There are no households that are included in both VHLSS and VLSS series. Further, VHLSS series before 2008 and after 2010 have no overlapping households. Because the survey do not track the households that have migrated, the location of the residence is the same for the panel households.

We remove those households without the consumption measurement. After this procedure, the original VLSS 1993 and 1998 data sets contain about 4,800 and 6,000 households, respectively. VHLSS 2002 covers about 30,000 households. In each VHLSS round between 2004 and 2014, slightly over 9,000 households are covered. For each of these data sets, about 20%–30% of households in northern rural Vietnam are used for Tables 1 and 9 as well as Figures 1–4.

Tables 2–8 and Figure 5 are based on panel components of these data sets. Panel VLSS 1993–1998 data contain 1,970 households or 3,940 observations. Panel VHLSS 2002–2004, 2004–2006,

2006–2008, 2010–2012 and 2012–2014 data contain, respectively, 1,728, 1,910, 1,865, 1,793 and 1,729 households.

APPENDIX B

SUMMARY OF DECOMPOSITION METHOD

In this section, we briefly summarise a variant of the regression-based decomposition method proposed by Fujii (2015). The only differences from Fujii (2015) are that the weight attached to each household is allowed to vary over time and the fixed-effects terms are explicitly included.

We start with the following individual-level poverty measure p_{ht}^e :

$$p_{ht}^e = g\left(\frac{y_{ht}^e}{z}\right) \mathbf{1}(y_{ht}^e \leq z) = g(\tilde{y}_{ht}^e) \mathbf{1}(\tilde{y}_{ht}^e \leq 1), \quad (2)$$

where z , $\tilde{y}_{ht}^e \equiv y_{ht}^e/z$ and $\mathbf{1}(\cdot)$ are, respectively, the poverty line, consumption per capita normalised by the poverty line and an indicator function, which takes one if the argument is true and zero otherwise. The function $g(\tilde{y}_{ht}^e)$ is once-differentiable for all $\tilde{y} > 0$ and $g(1) = 0$, which is satisfied for most commonly used poverty measures except for the head count index. In the special case where the poverty measure of interest is the poverty gap measure, we have $g(\tilde{y}^e) = 1 - \tilde{y}^e$, which clearly satisfies $g(1) = 0$. The poverty severity and Watts measures correspond to the cases where $g(\tilde{y}^e)$ is equal to $(1 - \tilde{y}^e)^2$ and $-\ln \tilde{y}^e$, respectively.

We consider the following additively decomposable measure of poverty P_t^e for the ethnic group e at time t , which can be calculated from p_{ht}^e for sample S^e :

$$P_t^e = \sum_{h \in S^e} w_{ht}^e p_{ht}^e,$$

where w_{ht}^e is the weight attached to household h at time t and satisfy $\sum_h w_{ht}^e = 1$ for all $t \in [0, 1]$ and $e \in \{0, 1\}$. In the case where there is no attrition and the sample is self-weighted, w_{ht} is constant over time and equal to the reciprocal of the sample size.

Without loss of generality, we assume that $t = 0$ is the initial time and $t = 1$ the terminal time for decomposition. Therefore, using the notations introduced above, the observed change in poverty $D^e \equiv P_1^e - P_0^e$ is ascribed to various components of interest in the decomposition analysis. Fujii (2015) derives the decomposition by the following transformation:

$$\begin{aligned} D^e &\equiv \sum_{h \in S^e} w_{h1}^e p_{h1}^e - w_{h0}^e p_{h0}^e = \sum_{h \in S^e} \int_0^1 \left[\frac{d}{dt} w_{ht}^e p_{ht}^e \right] dt = \sum_{h \in S^e} \int_0^1 (w_{ht}^e \dot{p}_{ht}^e + \dot{w}_{ht}^e p_{ht}^e) dt \\ &= \sum_{h \in S^e} \int_0^1 \left[w_{ht}^e g'(\tilde{y}_{ht}^e) \tilde{y}_{ht}^e \left(\dot{x}_{ht}^e \beta_t^e + x_{ht}^e \dot{\beta}_t^e + \dot{v}_{p(h)}^e + \dot{\epsilon}_{ht}^e \right) + \dot{w}_{ht}^e g(\tilde{y}_{ht}^e) \right] \mathbf{1}(\tilde{y}_{ht}^e \leq 1) dt, \end{aligned}$$

where the dot notation is used to denote the time derivative (e.g., $\dot{\beta} \equiv d\beta/dt$). We hereafter treat the provincial-level indicator variables as covariates and regard $v_{p(h)}^e$ as coefficients on them. The equation above shows that the change in poverty at the individual level can be decomposed into

the following four major components: (i) the covariate component (involving \dot{x}_{ht}^e), or the change in poverty due to the change in the distribution of covariates; (ii) the structural component (involving $\dot{\beta}_t^e$), or the change in poverty due to the change in the parameter β ; (iii) the weight component (involving \dot{w}_{ht}^e), or the change in the relative importance of households (due to the differences in fertility, mortality, and attrition across households); and (iv) the residual component (involving $\dot{\epsilon}_{ht}^e$), or the change in poverty due to the change in the error term. Notice that each of these components is additively decomposable because the poverty measure is additively decomposable.

One major practical issue that one faces when implementing this decomposition is that we would need a continuous observation of y_{ht}^e and x_{ht}^e . Therefore, we will need to make some assumptions about the underlying path along which x , w , β , and e changes. A reasonable assumption would be that each of these values change linearly over time between $t = 0$ and $t = 1$. Therefore, we make the following assumptions:

$$\begin{aligned}x_{ht,j}^e &= (1-t)x_{h0,j}^e + tx_{h1,j}^e, \\w_{ht}^e &= (1-t)w_{h0}^e + tw_{h1}^e, \\\beta_{t,j}^e &= (1-t)\beta_{0,j}^e + t\beta_{1,j}^e, \\\epsilon_{ht}^e &= (1-t)\epsilon_{h0}^e + t\epsilon_{h1}^e,\end{aligned}$$

where the subscript j denotes the j th component in the vector. To maintain the internal consistency, we apply the assumption for $x_{ht,j}^e$ only to the base term when the squared term is included. Therefore, our assumption implies that $[x_{ht,j}^e]^2 = [(1-t)x_{h0,j}^e + tx_{h1,j}^e]^2$ and not $[x_{ht,j}^e]^2 = (1-t)[x_{h0,j}^e]^2 + t[x_{h1,j}^e]^2$. A minor adjustment for X_j^e defined below is also made to reflect this point. Also, note that we are in effect assuming $v_{pt}^e = (1-t)v_{p0}^e + tv_{p1}^e$ because the term involving \dot{v}_{pt}^e is treated as a part of the structural component.

The covariates at the initial and terminal time points, x_{h0}^e and x_{h1}^e , are observed in the data and so are the weights w_{h0}^e and w_{h1}^e . The coefficients β_0^e and β_1^e and fixed-effects terms v_{p0}^e and v_{p1}^e can be estimated from the sample for the initial and terminal time points, respectively. Using these, the idiosyncratic error terms e_{h0}^e and e_{h1}^e can be also estimated.

Therefore, under these assumptions, we can obtain the following decomposition:

$$D^e = \sum_j (X_j^e + S_j^e) + W^e + R^e,$$

where X_j^e , S_j^e , W^e , and R^e are, respectively, the j th covariate component, j th structural component, weight component and residual component with the following definitions:

$$\begin{aligned}X_j^e &= \int_0^1 \sum_{i \in \mathcal{S}} w_{ht}^e g'(\tilde{y}_{ht}^e) \tilde{y}_{ht}^e (x_{i1,j}^e - x_{i0,j}^e) \beta_{t,j}^e \mathbf{1}(\tilde{y}_{ht}^e \leq 1) dt, \\S_j^e &= \int_0^1 \sum_{i \in \mathcal{S}} w_{ht}^e g'(\tilde{y}_{ht}^e) \tilde{y}_{ht}^e x_{it,j}^e (\beta_{i1,j}^e - \beta_{i0,j}^e) \mathbf{1}(\tilde{y}_{ht}^e \leq 1) dt, \\W^e &= \int_0^1 \sum_{i \in \mathcal{S}} (w_{i1} - w_{i0}) g(\tilde{y}_{ht}^e) \mathbf{1}(\tilde{y}_{ht}^e \leq 1) dt, \\R^e &= \frac{1}{n} \int_0^1 \sum_{i \in \mathcal{S}} w_{ht}^e g'(\tilde{y}_{ht}^e) \tilde{y}_{ht}^e (\ln \tilde{y}_{i1} - x_{i1}^e \beta_1^e - (\ln \tilde{y}_{i0} - x_{i0}^e \beta_0^e)) \mathbf{1}(\tilde{y}_{ht}^e \leq 1) dt.\end{aligned}$$

The integrals in the equations above are calculated using the fourth-order Runge–Kutta method. Each component is calculated separately for the ethnic minority and majority groups. We report the difference in each factor between the ethnic minority and majority groups in Table 7, where covariate and structural components are aggregated further to demographic, education, employment, land and provincial fixed-effects components for simplicity of presentation.