Climate Change and Intersectoral Labor Reallocation in the Presence of Labor Market Frictions[*]

Trinh Pham[†]

This Version: December 2022

Click Here for the Latest Version

Abstract

I study the impact of climate change on labor allocation in Vietnam by exploiting plausibly random year-to-year and decade-to-decade variations in weather distributions. Hot temperatures induce workers to reallocate from agriculture to non-agriculture in both the short and long terms. Reallocation rates, however, differ across age groups depending on destination jobs. While older individuals are likely to move into informal non-agriculture, younger workers comprise most of those who shift to a formal non-agricultural job. Supporting evidence suggests that these results are driven by relative labor productivity loss and labor market frictions, where workers move towards less affected sectors entry into which entails lower switching costs.

Keywords: Climate change, temperature, labor allocation, informality, labor market frictions, Vietnam

JEL Codes: O12, O17, Q54, J22, J24

[*]This paper has benefited from helpful comments and suggestions from Chris Barrett, Maulik Jagnani, John Hoddinott, Brian McCaig, and Ariel Ortiz-Bobea. I am grateful for feedback from seminar participants at Cornell University. All errors are mine.

[†]Dyson School of Applied Economics and Management, Cornell University. Email: tp347@cornell.edu; 410 Warren Hall, 137 Reservoir Ave, Ithaca, NY 14853, United States.

1 Introduction

Does temperature change accelerate or deter the reallocation of labor from agriculture to non-agricultural sectors, which Kuznets (1973) identifies as one of the six characteristics of modern economic growth? Theoretical models suggest at least two channels via which climate change can affect sectoral labor allocation: relative price effects and income effects (Kongsamut, Rebelo, and Xie 2001; Ngai and Pissarides 2007).[1] In an open economy where tradable prices are fixed by the world market, climate change only affects labor productivity, income, and thus labor supply. If the climate impact on labor productivity loss is larger in agriculture, such a differential change in relative labor productivity can push workers out of agriculture into other sectors of the economy. On the other hand, additional changes in local demand effects arising in general equilibrium might cause a reduction in demand for non-agricultural goods and employment. Likewise, when prices of goods and services are also endogenous, climate change can lead to an increase in the relative price of agricultural products, which can induce an increase in employment share for this sector.

These predictions are supported by previous empirical work, which documents differential temperature impacts on intersectoral labor reallocation. Under negative agricultural productivity growth caused by short-term increases in temperature, Colmer (2021) finds workers move away from agriculture to join manufacturing and services sectors within Indian labor markets. Studying the same context, Liu, Shamdasani, and Taraz (forthcoming) show, however, that local demand effects appear more significant in causing an increase in the agricultural labor share responding to rising temperature in the long run. Utilizing local short-term effects of temperature on agricultural yields and non-agricultural labor productivity, model simulations at the global scale by Nath (2020) also suggest that intersectoral labor reallocation is driven by local demand effects more than comparative advantage (i.e., climate change induces an inflow of workers to agriculture), and reducing trade barriers could substantially mitigate climate change damages.

In this paper, I seek to answer the following questions: does temperature change have different effects on sectoral labor allocation in the short and long terms? and does climate-employment relationship vary across demographic groups who incur different costs of movement depending on the type of job they take? A key fea-

[1] See Herrendorf, Rogerson, and Valentinyi (2014) for a review of the relevant literature.

ture of the empirical analysis is the utilization of both short-run weather variation and long-run climate variation to study the relationship between climate change and labor allocation among agricultural, formal and informal non-agricultural sectors for different age groups, thereby illustrating the role of labor market frictions, particularly switching costs, in the climate-employment relationship.[2]

While this paper does not aim to explain the structural sources of these frictions nor to identify remedies for them, its focus on employment in three sectors–agriculture, informal non-agriculture, and formal non-agriculture–makes it possible to infer the different extent of frictions that prevent free movement of workers and particularly of agricultural workers to other sectors within a simple framework and few assumptions. The focus of this paper is Vietnam, a low-middle income country that has experienced rapid structural transformation and growth with expanding informal and formal non-agricultural sectors over a period of relatively rapid warming that varies across sub-national regions (McCaig and Pavcnik 2015, 2017; Liu et al. 2020; World Bank n.d.).[3] The rate of change in sectoral employment shares, however, differs across age groups. The increase in formal non-agricultural employment share largely follows younger birth cohorts entering the labor market more into this sector than did prior cohorts of workers at those ages. For informal non-agriculture, in contrast, the change in employment share is largely due to economy-wide trends in which individuals of all birth cohorts move into this sector over time.

To examine the potentially differential effects of climate change on labor allocation by age group in a local market, I assemble a province-age group longitudinal dataset spanning nearly three decades from 1992 to 2018, linking measures of employment shares and hours worked in each sector for each age group with weather variables constructed using daily gridded weather data. I adopt two empirical specifications. The first approach exploits year-to-year variation in weather, while controlling for analysis unit fixed effects. This approach follows recent cli-

[2] I do not distinguish between informal and formal agriculture because agricultural production in developing countries is predominantly carried out by smallholder farmers, who generally do not have access to formal labor contract/social security, nor do they register with the government. Lowder, Skoet, and Raney (2016) estimate that at least 90% of farms in the world are held by individuals and families. According to the household surveys, in Vietnam, agricultural workers in formal firms account for approximately 2% of total agricultural employment over the study period.

[3] The rate of warming in the country is almost twice the global rate over the period 1971-2010 (World Bank n.d.). Vietnam has a diverse topography, long latitude, and is influenced by the East Sea, resulting in quite different climatic conditions across space.

mate impact literature (Dell, Jones, and Olken 2014) and relies on the identification assumption that conditional on province-by-age group fixed effects, and region-by-year fixed effects that vary across age groups, the variations in weather at the local level are orthogonal to unobserved determinants of sectoral employment in each province-age group cell. The second approach follows Burke and Emerick (2016) and exploits change in province-level temperature distributions over an extended period of up to 10-15 years, while controlling for region-by-age group trends. The inclusion of region fixed effects that vary across age groups in both approaches helps alleviate concerns over the potential conflation of an education effect with a temperature effect, where the former is correlated with age. It also addresses concerns about the non-monetary value of working in non-agriculture, especially formal non-agriculture, which may change differentially such that at any given point in time, the group of younger workers is less likely to work in agriculture when temperature increases anyway.

I capture in the empirical models different atmospheric elements of climate, including temperature, precipitation and wind speeds. In addition to common measures of temperature such as cumulative extreme temperatures (degree days), I also proxy for change in temperature distribution over a longer time frame using a statistical distance measure: the Kullback–Leibler divergence (KLD, henceforth). Because climate change involves not only changes in the mean but also in the variability of its properties and that persists for an extended period (IPCC 2022), the use of the KLD measure captures the change in the entire temperature distribution at the province level, thereby allowing an assessment of the relative importance of general warming–a rightward shift of the entire distribution–versus an increased risk of extreme temperatures in the climate-employment relationship. Throughout the paper, the analysis uses wet-bulb temperature to capture the combined effect of heat and humidity, which has been shown to be increasingly frequent across the globe due to climate change (Raymond, Matthews, and Horton 2020).[4]

Two main findings emerge about the relationship between climate change and labor allocation in Vietnam. First, rising temperatures accelerate the reallocation of labor away from agriculture to formal and informal non-agricultural sectors both in the short and long terms, and these effects entirely happen at the higher end of the temperature distribution.[5] While cold temperatures virtually do not affect

[4]The results are qualitatively similar when using dry-bulb temperature (air temperature).

[5]In general, the effects of other weather variables are qualitatively similar to that of hot temper-

sectoral employment shares, hot temperatures decrease the agricultural labor share and increase the formal and informal non-agricultural shares.

The magnitude of the estimated effect is economically meaningful. Estimations from the year-to-year panel specification show that every additional degree day with average wet-bulb temperature above 27°C–the 97th percentile of the historical temperature distribution–leads to a reduction of 0.6 percentage points in provincial-level agricultural employment shares, and increases of 0.38 and 0.27 percentage points in informal and formal non-agricultural employment shares, respectively. In contrast, there is no evidence of a temperature effect on the share of inactive and unemployed individuals.[6]

The effects of temperature on sectoral labor shares are consistent when being examined over a longer time frame. The increased risk of extreme temperatures, as opposed to a general warming, leads to a reduction in agricultural employment share and increases in both formal and informal non-agricultural employment shares. In addition, while the effect of long-term changes in hot temperatures remains similar in magnitude relative to the short-term effect for informal non-agricultural employment share, it intensifies for the share of workers in formal non-agriculture (and agriculture).

Second, these average effects mask significant heterogeneity by age group depending on the type of work climate change-induced sectoral migrants take. While older workers are as likely to move out of agriculture and into informal non-agriculture in response to hot temperatures, only the group of younger workers can take formal non-agricultural jobs. Again, these heterogeneous effects hold both in the short and long terms.

These findings are robust to different sample restrictions, alternative specifications (e.g., including time-variant demographic characteristics including education, share of ethnic minority population, share of male population, as well as to the inclusion of lagged dependent variables), to different methods of constructing weather variables (e.g., weighted average of the four closest grid points using inverse distance weighting, averaging values of grid points over a geographical

atures, for example, extremely high and low rainfall or episodes of high winds lead to a reduction in the agricultural labor share and corresponding increases in non-agricultural employment shares. These effects, however, are less robust to alternative specifications and thus are omitted from the discussion.

[6] The data do not allow me to distinguish whether an individual who did not work in the reference period was inactive in the labor market, or was unemployed.

boundary), and to different weather exposures. Results from the year-to-year panel approach are also robust to different functional forms of temperature, including fourth-order polynomials and cumulative degree day bins. Results from the long differences approach are generally robust to different period definitions.

I find evidence consistent with these results being driven by relative sectoral labor productivity loss and labor market frictions. Specifically, the positive (negative) effects of hot temperatures on non-agricultural (agricultural) employment shares are concentrated in reasonably tradable areas, as proxied by close distances to the major seaports of the country and a high correlation between local rice price with the world market price. Additional analyses show that hot temperatures, on average, significantly reduce labor supply and labor productivity in agriculture, but not in other sectors of the economy.[7] Taken together, this implies that when prices are fixed, the change in relative sectoral labor productivity and thus comparative advantage pushes workers to take up jobs in sectors that are less affected by temperature changes, which is consistent with the classic prediction from the structural transformation model of small open economies.

Using the Roy-Borjas model (Roy 1951; Borjas 1987) as a guiding framework, I further argue that the heterogeneous temperature effects by age group and sector of in-migration appear consistent with the existence of non-uniform labor market frictions, particularly switching costs that vary across sectors and age groups. Specifically, using a longitudinal dataset of workers who switched sectors, I show that the cost of switching from agriculture to informal non-agriculture is similar across age groups. However, transitioning into formal non-agriculture is significantly more costly for older workers relative to younger ones. As a result, given the relative labor productivity change induced by hot temperatures, workers of different age groups have similar likelihood of getting an informal non-agricultural job, but younger workers are much more likely to take up a job in formal non-agriculture.

This paper makes three contributions. First, I provide evidence on the consistent effects of temperature change, particularly at the right end of the distribution, on reallocation of workers away from agriculture to both informal and formal non-agricultural sectors in both short and long runs in the context of a developing country. This distinguishes the current work from previous studies, which generally

[7] It is noted that labor productivity of firms in climate-highly-exposed non-agricultural industries like mining and construction are also negatively affected by hot temperatures.

rely on short-term weather variation and/or focus on formal non-agricultural employment (Colmer 2021; Albert, Bustos, and Ponticelli 2021) or do not distinguish between work in informal and formal non-agricultural employment, or restrict the sample to only rural areas (Emerick 2018; Jessoe, Manning, and Taylor 2018; Liu, Shamdasani, and Taraz forthcoming).[8]

The results shed light on the role of relative labor productivity growth in studying structural transformation. Traditionally, scholars have argued that increasing agricultural productivity growth is an essential element of economic development.[9] In small open economies, however, the role of agricultural productivity can be very different. The reason for this is related to specialization forces according to comparative advantage. In this setting, I show that the disproportionately negative agricultural labor productivity growth caused by temperature changes is associated with an outflow of labor from agriculture and these effects are most pronounced in areas that are well-integrated into the world market.

Given that hot temperatures have detrimental effects on crop yields (Schlenker and Roberts 2009), earlier works on the climate and employment relationship often focus on temperature-induced agricultural productivity shocks as proxied by crop yields (Emerick 2018; Santangelo 2019; Colmer 2021; Liu, Shamdasani, and Taraz forthcoming). Here I find that the effect of hot temperatures on yields of rice–the country's main staple crop–is only one-third magnitude of the total effect of hot temperatures on annual revenue per agricultural worker, which could be partly attributed to a significant reduction in labor supply in response to hot temperatures. The results on disproportionate impacts of temperature on labor productivity in agriculture, as well as some other non-agricultural industries such as mining and quarrying, construction are consistent with earlier works that show larger temperature impacts on highly weather-exposed industries (Graff Zivin and Neidell 2014; Behrer and Park 2017). These findings support the hypothesis that climate change widens the pre-existing differences in intersectoral labor productivity, which could help drive structural transformation (Barrett, Ortiz-Bobea, and Pham forthcoming).

By separating informal and formal employment, these results raise another con-

[8] Exceptions include Colmer (2021), who shows a positive effect of short-run rising temperature on the number of workers in the informal manufacturing sector, although the effect is not statistically significant at conventional levels, and Liu, Shamdasani, and Taraz (forthcoming), who show an increase of non-agricultural labor shares in response to short-run increase in mean temperature, but an increase of agricultural employment share when examined over the long run.

[9] See, for example, Gollin (2014) for a review. Recent evidence by Bustos, Caprettini, and Ponticelli (2016) provides an illustration.

cern of whether such a reallocation would be welfare-enhancing to the total economy. While informality accounts for approximately 90% of the total employment and 70% of the non-agricultural employment in developing economies (Bonnet, Vanek, and Chen 2019), its role in economic development remains controversial. Informality offers flexible employment and is considered a free-entry sector of last resort on the one hand, but is associated with low productivity on the other hand (Fields 2009). Recent work further suggests that informality depresses human capital formation (Bobba et al. 2021), an important determinant of structural transformation (Porzio, Rossi, and Santangelo 2022). The fact that a large share of climate change-induced workers move into this sector might have important welfare consequences by reinforcing a country's comparative advantage in those less skill-intensive industries, which, if combined with low innovation, might lead to lower long run growth (Bustos et al. 2020).

Finally, this paper adds to the growing body of literature linking climate change and weather variables with inequality by showing a different aspect of the within-country inequality consequences of climate change. Previous literature on this topic primarily focuses on the effects of weather variables on various human capital and labor outcomes by gender, race or ethnicity (e.g., Maccini and Yang 2009; Park et al. 2020; Pham 2022). A common underlying mechanism of these results is the interaction of direct effects of climate anomalies and pre-existing gender bias in intrahousehold resource allocation, or differential access to coping and mitigation strategies due to socio-economic constraints. In this paper, I show that the temperature effects also vary across age cohorts and such effects are likely driven by frictions in the labor market. Specifically, mainly younger workers transition into the formal non-agricultural sector, whereas older workers either stay in agriculture, or take up an informal non-agricultural job in response to climate change. The large gaps in labor earnings across sectors, and especially the lack of social welfare system for workers in the informal economy suggest that these temperature effects may exacerbate the age gap in economic well-being in the country. This has important policy implications, given that the country has already begun its transition to an aged society (Glinskaya et al. 2021).

The rest of the paper proceeds as follows. I describe the data sources and key employment variables of interest in section 2. In section 3, I discuss the measure of temperature change and present descriptive patterns of the temperature change-sectoral employment share relationship over the last three decades. Section 4 details empirical approaches to estimate causal effects of temperature change on sectoral labor allocation, discusses the main findings, as well as reports results from a series of robustness checks and placebo tests. Section 5 explores potential mechanisms underlying the main results and section 6 concludes the paper.

2 Data and Measurement

In this section, I briefly discuss the main data sources and variables of interest. For detailed variable definition and data construction, see Appendix A. The main sources of data include the Vietnamese household surveys, the population census, and the global climate and weather reanalysis ERA5 database.

2.1 Employment Data

I use the 5% random sample of the 1989 population census, the Vietnam Living Standards Surveys 1993-1998, the Vietnam Household Living Standards Surveys 2002-2018 to construct measures of the sectoral composition of employment and sectoral hours of work. The surveys are nationally and provincially representative.[10] Although the household survey is a repeated cross-sectional survey, it contains a rotating panel sub-component that tracks individuals over a period of up to four years, which allows me to analyze individual transition across sectors over a longer time than is usually feasible.[11] The analysis sample includes workers with available information on industry of employment, as well as types of employers for household members age 24-64. I focus on this age range to capture working-age individuals with completed education.

The key variable of interest is the sector in which an individual was working

[10] The surveys use households as sampling units and define household membership on the basis of physical presence. Individuals must stay and eat in the household for at least six months during the 12-month reference period, and contribute to collective income and expenses to be considered members. This requirement means that individuals who have moved away for work or school (e.g., migrants) are not considered household members. Considering an individual as a seasonal migrant if they left the household for work during the year but are still considered as a household member (as in Brauw and Harigaya (2007)), then more than 96% of household members staying in households during the last 12 months also suggests a low seasonal migration rate of 4%.

[11] More recent effort has been made to collect longitudinal individual data to study employment transitions in developing countries, for example, in Indonesia and Kenya (Hamory et al. 2021). In documenting the relationship between labor market dynamics and economic development, Donovan, Lu, and Schoellman (2021) construct a dataset of gross labor market flows of individuals of up to 6-9 months for a sample of 45 countries.

during the reference period and their working hours. This variable is constructed using data from the employment modules of the census, which covers industries of the most time-consuming job, as well as of the household survey, which covers hours worked, industries, and types of employer of the two most time-consuming jobs.[12] For each job, an individual is asked whether he or she works for his or her own household or for other households, collectives, state-owned enterprises, private domestic enterprises, or foreign-invested enterprises. Following McCaig and Pavcnik (2015, 2018), I consider an individual as working in the informal sector if he or she is self-employed or works as an employee in household businesses. I also consider working in collectives or cooperatives as informal in order to make the definition consistent over a longer time period, although this should not affect the analysis much.[13]

Note that informality can broadly be defined either from the worker side or from the employer side. According to GSO and ILO (2018), informality on the worker side implies that workers do not have social security benefits and a labor contract with a minimum term of three months ("informal workers"). On the employer side, informality implies that firms do not register with the government ("informal firms"). A cross-check whenever possible suggests that the notion of informality employed in this paper is highly correlated with the definition of informality from the worker side, with a Pearson correlation coefficient of approximately 0.9. However, while only a small fraction of formal workers labor in informal firms, a nontrivial 14% of workers in formal firms does not have social security benefits and a labor contract and most of them work in medium tech manufacturing industries, less knowledge-intensive service industries, mining and quarrying, and construc-

[12]Data on secondary job are not available in 2002. Since 2010, VHLSS asks if the individual also works a third job for wage. Approximately 2.7% of the working population answered yes to this question. For a majority of these workers (75%), agriculture is their primary sector, followed by construction. Information regarding hours worked, earnings, and industries are not available for the third job and beyond.

[13]Before the "Doi Moi" reform in 1986, Vietnamese economy was centrally-planned and there was no market-based price mechanism. Without an enterprise law, all industrial producers and traders were owned by the government. Agriculture was required by the state to operate in the form of village-level collectives (Nguyen, Luu, and Trinh 2016). Since the late 1980s and early 1990s, however, the formation of collectives has been voluntary with households essentially exchanging labor during plowing, planting, and harvesting seasons (Raymond 2008). Furthermore, while it is not officially stated in the first Enterprise Law enacted in 2000, the Cooperative Law of 2012 emphasizes that a collective or a cooperative is not considered a type of enterprise. As such, the notion of collectives resembles that of household businesses, which is the main source of informal employment used by McCaig and Pavcnik (2015, 2018). Employment in cooperatives and collectives since 2000 contributes to less than 1% of the total employment of adults 24-64 years old.

tion.[14] If climate change induces reallocation of workers into temporary jobs, as in Colmer (2021), then the fact that a nontrivial share of workers in formal firms are informally employed suggests this paper likely underestimates the role of the intensive margin of informality in the Vietnamese economy in response to climate change.[15]

2.2 Historical Weather Data

The main weather data are from ERA5 reanalysis, which combines model data with observations from across the world into a globally complete and consistent dataset (Hersbach et al. 2020) and contains hourly atmospheric variables for the period on 0.25°× 0.25°grid (approximately 30 km at the equator). Reanalysis data provide a consistent estimate of atmospheric parameters over time and space (Auffhammer et al. 2013), and have been increasingly used in the literature, especially in developing countries where the quality and quantity of weather data are limited (Ortiz-Bobea 2021). The variables I focus on are grid-level daily mean wet-bulb temperature, precipitation, and wind speed over the study period. The decision to use daily mean values of temperature instead of maxima stems from the fact that reanalysis data, which are outputs from climate model prediction, are generally sensitive to extreme values. While most reanalysis datasets agree on the mean value of weather variables across space, they are not in full agreement about the timing or magnitude of deviations from this mean (Auffhammer et al. 2013).[16]

Grid-level weather data are then transformed to province-level weather data by taking weighted average of the four nearest grid points to the geographic centroid of the first administrative level–a province, with weights being the inverse distance. Because there have been changes in administrative boundary in Vietnam over the study period, and most of the changes happens in the case of splitting, I use the original administrative units in 1993, which gives a consistent sample of 52 provinces over the study period.[17]

[14] For details on share of informal workers in formal firms by industry, see Appendix A2. Industries are ranked following the Statistical Classification of Economic Activities in the European Community. For details, see Annex 3 – High-tech Aggregation by Statistical Classification of Economic Activities in the European Community (NACE Rev.2)

[15] The notions of extensive and intensive margins of informality follow Ulyssea (2018).

[16] In settings where station-level data are of good quality and available at high spatio-temporal density, maximum or minimum values of weather variables are commonly used (e.g., Graff Zivin and Neidell (2014)).

[17] An exception is the then Ha Tay province, which was merged into Hanoi city in 2008, and there-

2.3 Other Data

In addition to employment and weather data, I construct a province-level panel dataset of labor productivity, as proxied by revenue per worker, in three sectors (agriculture, informal and formal non-agriculture) from the enterprise census and the household survey. I also compile a dataset of migration rates (in-, out-, and net-migration) and agricultural yields from statistical yearbook.

2.4 Merging Employment Data with Weather Data

The employment data are constructed on an individual-level basis. Employment variables including sector of employment and hours of work are recorded for the 12-month reference period prior to the interview day. Individuals from different households in the same province may not have same exposure to the weather distribution during their reference period because the survey is typically conducted in different months throughout the year for each province. Given that more than 96% of household members stay in the same province over the full reference period, I assume that individual j surveyed in month m of year t in province p has been exposed to the weather distribution of province p during the 12 full months prior to m.[18] Individual-year employment data and weather data are then collapsed to province-age group-year level (or province-year level, depending on the analysis) by computing the weighted mean, where weights are the survey sampling weights.

3 Descriptive Patterns of Climate Change and Employment

3.1 Measure of Climate Change

Climate change refers to an alteration of climate that persists for an extended period. While much of the public attention has been focused on the accelerating increase in global mean air temperature of about 1°C or so over the last four decades (Hsiang and Kopp 2018), there was also a doubling in the frequency of dangerous combinations of heat and humidity across the globe (Raymond, Matthews, and Horton 2020). One metric closely related to the combined effects of heat and humidity is wet-bulb temperature (WBT). At the same level of heat, a place with

fore I use the boundary of the new Ha Noi for consistency.

[18] For example, the individual j surveyed in January of 2010 in province p is assumed to be exposed to the weather distribution of province p from January to December of 2009.

dry air has a lower wet-bulb temperature and feels cooler compared to a place with humid air, because the former allows quicker evaporation of sweat in order to avoid overheating, the process that negatively affects human health and productivity. Climate models have consistently predicted an increase in WBT levels, which can exceed the 35°C "survival" threshold in some places including, but not limited to, the tropical regions (Sherwood and Huber 2010; Zhang, Held, and Fueglistaler 2021), where a substantial share of the global population lives in poor conditions with limited adaptation capacity (IPCC 2022).

Due to its diverse topography, long latitude, and influences from the East Sea, Vietnam is characterized by different climatic conditions that vary greatly between regions. According to Beck et al. (2018), the country can be classified into seven climatic regions including tropical-rainforest, tropical-monsoon, tropical-savannah, arid-steppe-hot, temperate-dry winter-hot summer, temperate-dry winter-warm summer, and temperate-no dry season-hot summer. These roughly correspond to the country's seven economic regions.

To capture the change in temperature distribution over time, I use the Kullback-Leibler divergence (KLD, henceforth), a measure of how one distribution differs from another. A KLD value of zero implies two distributions are identical, and a greater value implies more difference between the two distributions.[19] The difference between two distributions can be decomposed into two components, namely location and shape. Location difference arises when the distribution of daily temperature in the recent period differs from the distribution in the reference period because of a general (rightward) shift that affects all points along the distribution to the same extent. General warming would manifest as a significantly positive location difference. Shape difference, on the other hand, refers to a change in the structure (pattern) of the distribution conditional on location, for example, fewer mild days and more extreme days that lead to a more "polarization" of the temperature distribution in recent years. An increased risk of extreme temperatures would appear as a significantly positive shape difference.

Appendix Figure C1 illustrates the difference between location and shape components of two provinces that experience temperature rises with similar overall divergence in temperature distribution and increase in mean temperature but different extent of shape and location effects over the period 1992-2018. When assuming similar locations of the recent and reference distributions, the shape of the

[19] For a review of different measures and computation in Stata, see Jann (2021).

recent distribution is similar to the reference distribution in Lam Dong. In Binh Dinh, although less precisely estimated, there appears to be a significant change in the shape of the distribution with fewer mid-range days and and more days on the right tail.

Distinguishing location and shape differences is important to assess the effect of climate change. Although a change in either location or shape of the temperature distribution can lead to a similar increase in the mean, the latter is associated with more variation and thus is generally less predictable, leaving less room for adaptation. The ecology literature has emphasized greater risks and effects that changes in temperature variation, relative to changes in temperature mean, may pose to ecological systems (Vasseur et al. 2014; Turner et al. 2020). This is particularly concerning given that increased temperature variability has been consistently projected to be more prevalent in poor tropical countries in the near future (Bathiany et al. 2018).

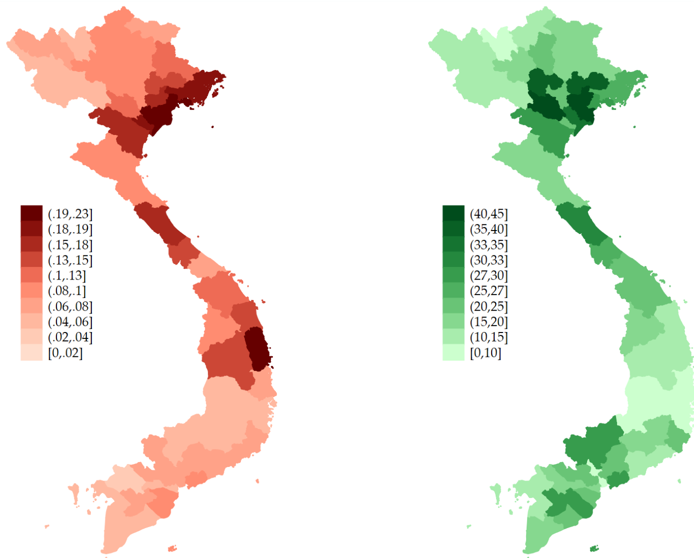
3.2 Association between Change in Temperature Distribution and Sectoral Employment Shares

Panel A of Figure 1 presents the change in temperature distribution in Vietnamese provinces over two periods: 1992-2006 and 2007-2018 using the measure of shape difference. The change in temperature is heterogeneous across the country, with the Red River Delta, central coast and the southeast regions experiencing the most change. Correspondingly, these regions also observe a relatively larger decrease in the average agricultural employment share between the two periods.

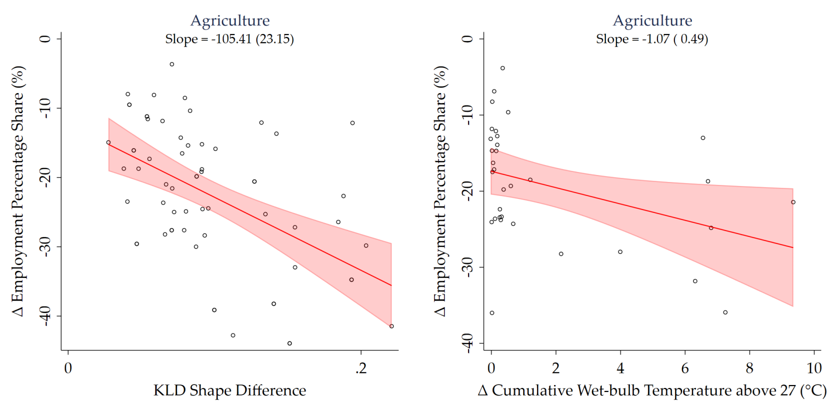
Panel B (left) confirms a clear negative association between the shape difference measure and the decrease in agricultural employment share at the province level. The right panel likewise plots the relationship between change in extreme hot temperatures, as proxied by cumulative wet-bulb temperature degree days above 27°C, and sectoral employment share changes. The temperature-employment relationship appears to be similar with the use of this conventional measure, suggesting that KLD shape difference captures relatively well the increased risk of extreme temperatures.

Figure 1: Change in Temperature Distribution and Sectoral Employment Shares: 1992-2018

Panel A: Shape Difference (Left) and Decrease in Agricultural Employment Share (Right)



Panel B: Relationship between Temperature Change and Agricultural Employment Share Change



NOTES: In panel A, darker color denotes larger change/value. Panel B shows the line of best fit and 95% confidence interval from the regression of change in sectoral employment shares on change in wet-bulb temperature distribution, as proxied by shape difference (left) and change in extreme temperatures (right), with each circle representing a province.

3.3 Association between Change in Temperature Distribution and Sectoral Employment Shares by Age Group

As observed in many other countries, a key feature of the evolution of sectoral employment shares in Vietnam is the stark difference in the rate of labor reallocation across age and birth cohorts. Figure 2 shows the share of workers in each sector for four 4-year intervals from 1989 to 2018 and for five 4-year age intervals. As seen, those ages 24-28 are 40% less likely to work in agriculture in 2014-2018 than people in that age range were in 1989-1993. The corresponding figure for the group of older workers (age 56-60) is only 20%. In terms of change within birth cohorts, the agricultural employment of those entering the market in the period 1989-1993 decreased from 68% (at ages 24-28) to 46% at ages 48-52. This life cycle pattern appears to hold through the 1998-2002 entrant cohort. Younger cohorts also enter the labor market more in the formal non-agricultural sector. For informal non-agriculture, in contrast, the change in employment share largely follows economy-wide trends in which individuals of all birth cohorts move into this sector over time.[20]

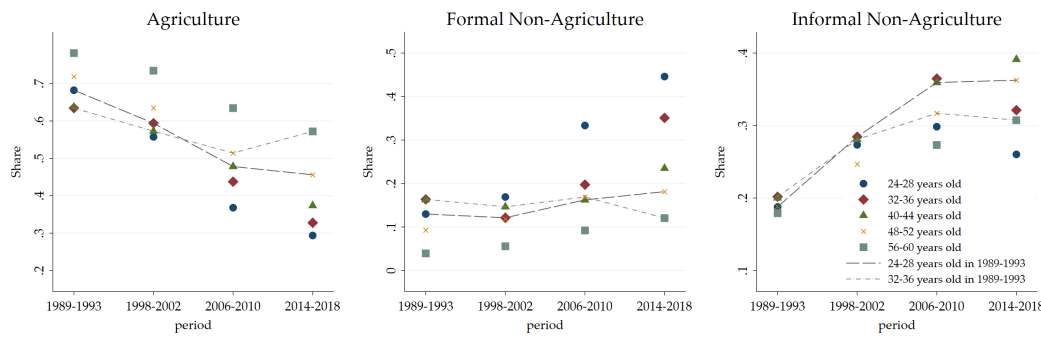
Appendix Figure C2 plots the relationship between changes in temperature distribution and changes in sectoral employment share for three age groups: 24-39, 40-54, and 55-64. The temperature change-employment share relationship patterns mirror that of the nationwide changes in sectoral employment shares by age group. In particular, the association between temperature change and agricultural or formal non-agricultural labor share is stronger for the younger group. In contrast, the positive relationship between temperature and informal non-agricultural employment share appears similar across groups.

4 The Causal Effect of Temperature on Sectoral Employment

Section 3 provided some descriptive evidence that there is a relationship between within-province change in temperature distribution and within-province change in

[20]These patterns are also observed by Kim and Topel (1995) and Hobijn, Schoellman, and Vindas (2019), who show the important role of across-cohort reallocation in declining agricultural employment shares in South Korea and the US, respectively. They are also consistent with Porzio, Rossi, and Santangelo (2022), who extend the analysis to a larger set of countries across the global income distribution and show that the particular role of new cohorts in labor reallocation out of agriculture ties to the expansion of education that equips younger cohorts with skills more valued outside of agriculture.

Figure 2: Sectoral Employment Shares by Age Group and Year



NOTES: This figure plots the share of workers in each sector for four 4-year intervals from 1989 to 2018 and for five 4-year age intervals. While change in agricultural and formal employment share largely follows younger cohorts entering the labor market more into this sector, the change in informal employment share is largely due to economy-wide trends in which individuals of all birth cohorts move into this sector over time.

sectoral employment shares and such a relationship varies across age groups. Such observations, however, could have been driven by a change in frequency and intensity of extreme weather events, or change in demographic characteristics, including educational attainment, that are correlated over time with rising temperatures. To examine the causal effect of temperature change on sectoral labor allocation, I apply two different approaches, each of which relies on a different identification strategy and source of variation in the weather distribution.

4.1 Empirical Strategy

4.1.1 Panel Approach

The first approach exploits year-to-year variation in weather distribution within geographic and demographic cells. I estimate the regression of the form:

$$y_{aprt} = f(a, \mathrm{WBT}_{pt}) + g(a, \mathrm{R}_{pt}) + \gamma_{ap} + \gamma_{art} + \varepsilon_{apt} \tag{1}$$

where y_{aprt} is an outcome of age group $a \in \{24-39, 40-54, 55-64\}$ of province p in region r in year t. The outcomes include employment shares in agriculture, informal non-agriculture, and formal non-agriculture for the main job. The term R_{pt} represents a vector of weather variables other than temperature in province p in the reference period relative to year t, including second-degree polynomials of rainfall and episodes of high speed wind, which is allowed to have differential effects by age group.[21]

Equation (1) includes a full set of province-by-age group fixed effects γ_{ap}, which absorb all unobserved, province-specific time-invariant determinants of sectoral employment for each age group. The equation also includes region-by-year fixed effects that are allowed to vary across the age groups γ_{art}. These fixed effects control for time-varying differences in the dependent variable that are common across provinces within age groups in a region, for example, regional economic development policy aims to boost industrial sector will generate demand for (formal) non-agricultural labor, especially young workers, which might lead to an increase in non-agricultural employment shares in the absence of temperature change. Across all estimations with equation (1), I weight the results by age group-specific population so that the coefficients correspond to an average person in the relevant age category.[22] I cluster standard errors at the province level to allow for potential serial correlation over time within each province. I also report Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags in the error term (Conley 1999).[23]

The focus of equation (1) is on the effect of temperature on sectoral employment, represented by the response function $f(a, \text{WBT}_{pt})$ that varies by age group. In the most parsimonious model, I define $f(a, \text{WBT}_{pt})$ as a piece-wise linear function:

$$f(a, \text{WBT}_{pt}) = \begin{cases} \sum_a \sum_{d=1}^{365} \beta_{a11}(11 - \text{WBT}_{dpt})\mathbb{I}_a & \text{if} \quad 0 \leq \text{WBT} < 11 \\ 0 & \text{if} \quad 11 \leq \text{WBT} < 27 \\ \sum_a \sum_{d=1}^{365} \beta_{a27}(\text{WBT}_{dpt} - 27)\mathbb{I}_a & \text{if} \quad \text{WBT} \geq 27 \end{cases} \tag{2}$$

With this function, β_{a11} and β_{a27} can be interpreted as the effect of one additional degree day below 11°C or above 27°C, respectively, on sectoral employment shares

[21] A day is considered having high speed wind if its maximum wind speed is above 10.8 m/s, which corresponds to the Beaufort scale level 6 (strong breeze). According to the INFORM risk index database, the country has been facing high natural hazard risks such as floods, followed by cyclones.

[22] Weights are constructed based on survey sampling weights so that they sum to one for each survey year in the sample, across all observations. Specifically, the weight for an observation of age group a in province p in year t is $w_{apt} = \frac{\sum_{i \in a} sw_{ipt}}{\sum_p sw_{ipt}}$ where sw_{ipt} is the sampling weights for each individual observation i available in the survey for year t.

[23] The choice of five lags is arbitrary. Results are robust to other choices of lags. I implement Conley standard errors in Stata using the module that allows weighting developed by Colella et al. (2019).

of the age group a during the 12-month reference period.[24] These cutoffs roughly correspond to the bottom and top 3% of the daily wet-bulb temperature distribution in the sample over the study period (Appendix Figure C3).

Theoretically, predictions for the signs of β_{a11} and β_{a27} are ambiguous. If cold and hot temperatures cause greater productivity loss in agriculture relative to informal and formal non-agriculture then they would lead to a decrease in agricultural employment shares. Under this scenario, the sign of the coefficients on temperatures will be negative for the agriculture specification and positive for the non-agriculture specifications. On the other hand, if cold and hot temperatures affect household income and additionally generate the "food problem,"–where expenditure share on agricultural output increases relative to non-agricultural goods and services–then the local demand effects may dominate, with increasing (decreasing) agricultural (non-agricultural) employment shares. In this case, the corresponding coefficients will be of opposite directions to the previous scenario. Likewise, predictions for the relative magnitude of β_{11} and β_{27} across age groups are ex-ante ambiguous.

As a robustness check, I also estimate a model where $f(.)$ is represented by cumulative temperature bins, degree day bins, as well as fourth-order polynomials of daily average temperatures, summed across year.[25] As for cumulative temperature bins and degree day bins, denote the endpoints of the eleven temperature bins (less than 9°C, 9 two-degree wide bins, higher than 27°C) by $[\text{WBT}^{1k}, \text{WBT}^{2k})$ with $k \in \{1, 2, ..., 11\}$. For degree day bins, I follow Somanathan et al. (2021) and consider a daily mean WBT contributes positive degrees to the bin for which $\text{WBT} > \text{WBT}^{1k}$ and zero to all others. If $\text{WBT} \geq \text{WBT}^{2k}$, the day contributes $\text{WBT}^{2k} - \text{WBT}^{1k}$ to bin k; if $\text{WBT}^{1k} < \text{WBT} \leq \text{WBT}^{2k}$ then it contributes $\text{WBT} - \text{WBT}^{1k}$ to bin k. In the cumulative temperature bins, a day with WBT contributes positive degrees to the bin for which $\text{WBT}^{1k} < \text{WBT} \leq \text{WBT}^{2k}$ and zero to all others. I assume that for $\text{WBT} < 19$°C, the day contributes $\text{WBT}^{2k} - \text{WBT}$ to bin k; while for $\text{WBT} > 19$°C,

[24] For example, if a given province-year-age group experienced two days over 27°C, one at 28°C and the other at 30°C, its value of degree day above 27°C would be 4 (i.e., 1 + 3).

[25] Following Carleton et al. (2022), in order to preserve the non-linear relationship between weather variables and sectoral employment share that occurs at the grid cell level, although the equation (1) is estimated at a higher level of aggregation, I first raise grid-level daily weather variables to the power $n \in \{1, 2, 3, 4\}$, then take a weighted average of these values of the four grid points nearest to the geographical centroid of province p, where the weight is inverse distance. I then sum these daily polynomial terms over days during the reference period of individual i in province p before collapsing into province-age group-year cell.

the day contributes $\text{WBT} - \text{WBT}^{1k}$ to bin k. These values are then summed across the reference period to determine the number of degrees in each bin.[26] These models provide sufficient flexibility to capture important non-linearity, as well as being relatively parsimonious with low demand on the data.

Identification Assumption The validity of estimates based on equation (1) relies on the assumption that $\mathbb{E}\left[f\left(a, \text{WBT}_{pt}\right)\varepsilon_{aprt}|g(a,\text{R}), \gamma_{ap}, \gamma_{art}\right] = 0$. By conditioning on other weather variables, province-by age group fixed effects and region-by-age group-by-year fixed effects, these coefficients are identified from province-age group-specific deviations in temperature distribution about its averages after controlling for shocks that could affect different age groups of different regions to different extent. The inclusion of region-by-age group-by-year fixed effects is to address the case in which, for example, the non-monetary value of working in non-agriculture, especially formal non-agriculture, might grow differently such that at any given point in time, the younger cohorts are less likely working in agriculture in regions when temperature increases anyway. This also addresses the concern that the common trend in both educational attainment and in temperature, where the former is correlated with age, might conflate an education effect with a temperature effect. The identifying variation is assumed to be orthogonal to unobserved determinants of sectoral employment in each age group-province cell.[27]

4.1.2 Long Differences Approach

The second approach exploits variation in the long-term change in temperature and thus the estimates can be interpreted as the long-run responses to climate change. I follow Burke and Emerick (2016) and estimate a long differences regression of the following form:

$$\Delta y_{apr} = f_{\text{LD}}(a, \Delta\text{WBT}_p) + g_{\text{LD}}(a, \Delta\text{R}_p) + \gamma_{ar} + \varepsilon_{apr} \quad (3)$$

[26] The interpretation of these cumulative temperature bin coefficients therefore is similar to the baseline parsimonious model, for example, the coefficient on the bin [25 – 27) represents the effect of one additional degree day with WBT higher than 25°C (but lower than 27°C), and the coefficient on the bin [15 – 17) denotes one additional degree day with WBT lower than 17°C (but higher than 15°C).

[27] In Appendix B, I explore the robustness of the results to inclusion of time-varying demographic characteristics, including educational attainment.

where Δy_{apr} represents the change in sectoral employment shares of age group a of province p between the two periods 1992-2006 and 2007-2018.[28] The shares in each period are calculated as the average of the shares in each survey waves during that period.[29] The term $\Delta \mathrm{WBT}_p$ denotes change in wet-bulb temperature distribution. As discussed, the effect of temperature change might be conflated with that of precipitation or other weather events, which I address by including in $\Delta \mathrm{R}_p$ change in episodes of extreme high/low precipitation relative to long-term mean where the long-term mean is determined over the period 1980-2020, as well as number of days with high wind speeds.[30] Equation (3) also includes region-by-age group fixed effects γ_{ar}, which controls for any unobserved trends at the climatic or economic region level that vary by age group. I report standard errors clustered at the province level, as well as Conley standard errors that allow for spatial correlation up to 200 km (Conley 1999).[31]

I estimate two specifications of equation (3). The first specification uses the measures of KLD location and shape differences, which are calculated using the calendar year temperature distribution of the periods 1992-2006 and 2007-2018, to proxy for $\Delta \mathrm{WBT}_p$. In other words, all individuals (and thus all age groups) in the same province are assumed to experience the same weather distribution, regardless of their interview timing. Because the KLD measures reflect the difference between two probability distributions, the decision to use calendar year instead of reference year according to interview timing is to correct for bias in distribution changes that are mechanically driven by changes in interview timing. As long as the average individual in each province shares similar interview timing, the coefficients on KLD shape and location are reflective of the averge effects of province-level temperature change on province-level sectoral employment shares. As seen in Appendix Figure C4, for most provinces, the average individual was surveyed between July-August, which supports this interpretation.

Although KLD measures are useful to proxy for change in the temperature dis-

[28] In Appendix B, I explore the robustness of the results to other splits of the time series for differencing.

[29] For example, the first period comprises of five waves of household surveys 1992/1993, 1998, 2002, 2004, and 2006.

[30] Excess precipitation is calculated in terms of monthly standardized precipitation index. Details in Appendix A1.2.

[31] I do not find any evidence of spatial correlation in residuals from the Moran test conducted after estimating equation (3) for each outcome (and age group) separately. As seen below, the two standard errors are pretty similar in magnitude. In some cases, the Conley standard errors are smaller than the corresponding ones clustered at the province level.

tribution over an extended period of time, they are not without drawbacks: the measure is difficult to interpret, and is not a metric.[32] To facilitate a comparison between the coefficients estimated using long-term change in climate and those estimated using short-term weather variation, I estimate a variant of equation (3), with the temperature response function being defined as:

$$f_{\text{LD}}(a, \Delta \text{WBT}_p) = \begin{cases} \sum_a \beta_{a11,\text{LD}} \Delta(11 - \text{WBT}_{pt}) \mathbb{I}_a & \text{if} \quad 0 \leq \text{WBT} < 11 \\ 0 & \text{if} \quad 11 \leq \text{WBT} < 27 \\ \sum_a \beta_{a27,\text{LD}} \Delta(\text{WBT}_{pt} - 27) \mathbb{I}_a & \text{if} \quad \text{WBT} \geq 27 \end{cases} \tag{4}$$

where $\mathbb{I}$ is an indicator function. In words, the function represents the effect of change in extreme temperatures: the difference in the average amount of degree days lower than 11°C and higher than 27°C wet-bulb temperature between any two periods. This temperature measure is constructed using the survey reference period, similar to the panel approach, and thus, the results can be directly compared.

4.2 Empirical Results

4.2.1 Panel Approach

Average Temperature Effects Table 1 reports the results from estimating a variant of equation (1) in which the response functions $f(.)$ and $g(.)$ are not specific to any age group. The estimates can be interpreted as the percentage point change in sectoral employment share resulting from one additional degree day below 11°C or higher than 27°C wet-bulb temperature during the 12-month reference period. Columns (1)-(3) increase the saturation of temporal controls in the model specification: column (1) controls for year-specific unobserved common shocks that affect all age groups within each region to the same extent; column (2) adds age-specific linear trends, allowing different age groups to follow different trends in a limited but potentially revealing way. In column (3), which is the preferred specification, different age groups of different regions are assumed to follow more flexible trends.

Table 1 shows that cold temperatures do not affect sectoral labor shares. One extra day with wet-bulb temperature higher than 27°C, however, decreases the

[32] Specifically, KLD does not satisfy symmetry and triangle inequality (Amari 2016), which makes the year-to-year interpretation unwarranted.

Table 1: Wet-bulb Temperature and Primary Sectoral Employment

Panel Approach

	Agriculture			Formal Non-Agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)
(11 – WBT)	0.001	-0.000	-0.004	-0.014	-0.012	-0.015
	(0.037)	(0.037)	(0.039)	(0.015)	(0.015)	(0.016)
.	[0.039]	[0.041]	[0.043]	[0.019]	[0.021]	[0.021]
(WBT – 27)	-0.529	-0.566	-0.598	0.206	0.247	0.274
	(0.154)	(0.154)	(0.166)	(0.118)	(0.105)	(0.105)
.	[0.162]	[0.186]	[0.197]	[0.067]	[0.098]	[0.100]
Sample Mean of DepVar	45.01	45.01	45.01	17.10	17.10	17.10
	Informal Non-Agriculture			**No Employment**		
	(1)	(2)	(3)	(4)	(5)	(6)
(11 – WBT)	0.017	0.016	0.021	-0.009	-0.009	-0.008
	(0.030)	(0.030)	(0.031)	(0.006)	(0.006)	(0.006)
.	[0.028]	[0.028]	[0.030]	[0.006]	[0.006]	[0.007]
(WBT – 27)	0.381	0.371	0.381	-0.059	-0.052	-0.058
	(0.138)	(0.139)	(0.142)	(0.079)	(0.080)	(0.081)
.	[0.158]	[0.157]	[0.158]	[0.066]	[0.066]	[0.064]
Sample Mean of DepVar	27.69	27.69	27.69	8.65	8.65	8.65
Observations	1707	1707	1707	1707	1707	1707
Province × Age Group FE	x	x	x	x	x	x
Region × Year FE	x	x		x	x	
Age Group Linear Trend		x			x	
Region × Age Group × Year FE			x			x

NOTES: Unit of analysis is province-age group-year. Dependent variables are percentage shares of employment in each sector. All columns control for the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. All regressions use sampling weights.

provincial-level employment share in agriculture by roughly 0.6 percentage points (p-value < 0.01). The corresponding effects on formal and informal non-agricultural labor shares are 0.27 and 0.38 percentage points, respectively. These effects are statistically significant at the 5% level when standard errors are clustered at the province level and account for spatial and temporal correlation. To put these effects into perspective, an average person in the sample experiences approximately 4.7 cumulative degree days higher than the 27°C threshold during the 12-month reference period, so the change in agricultural labor share induced by hot temperatures amounts to approximately 6.2% of the sample mean. By the same argument, the increases in formal and informal non-agricultural employment shares induced by hot temperatures are roughly 7.5% and 6.4% of the corresponding sample means. Because there is no significant change in no employment, these findings do not reflect a labor supply effect, but a labor reallocation effect.

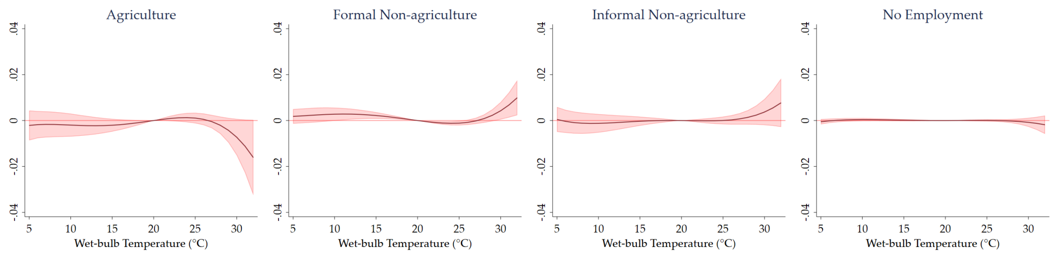
Figure 3 presents additional results with the temperature function being represented by fourth-order polynomials of daily average temperatures, cumulative temperature bins, and degree day bins. As seen, the results from these alternative models are consistent with those obtained from the model with extreme temperatures: all of the labor reallocation effects are driven by the higher end of the temperature distribution.

Heterogeneous Effects by Age Group Table 2 shows the effects of temperatures on three group of workers, which are estimated from equation (1).[33] As seen, younger workers are less likely to work in agriculture in response to hot temperatures: one extra degree day higher than 27°C wet-bulb temperature decreases agricultural labor share for workers age 24-39 by roughly 0.8 percentage points. The corresponding effect for workers age 40-54 is 0.5 percentage points. The effect on older workers is less precisely estimated, with the near-zero point estimate suggesting that they are virtually not affected by hot temperatures. Correspondingly, the two younger groups also experience significant increase in formal non-agricultural employment shares, with some suggestive evidence of the largest effect among the youngest group. Turning to informal non-agricultural employment, there is no statistical difference in the temperature effects across age groups.

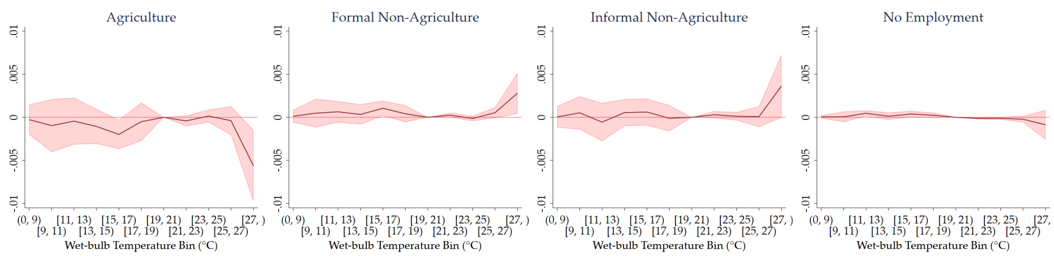
Piecing results across the three sectors, it appears that each age group has a different response to hot temperatures. While workers are more likely to leave agriculture as a consequence of temperature changes, younger workers are more

[33] For the sake of brevity, the similar and statistically insignificant effects of cold temperatures across age groups are omitted from the table.

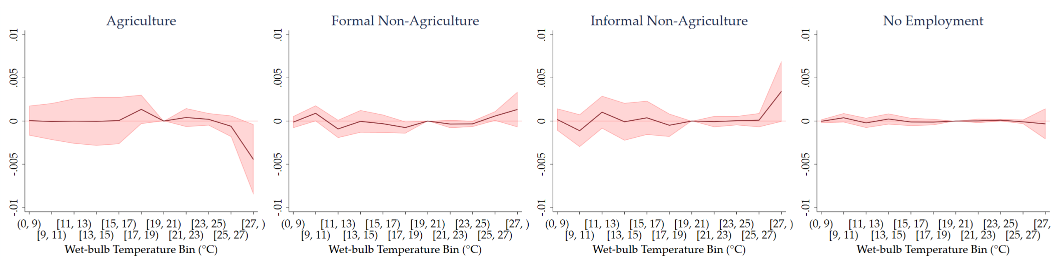
Figure 3: Robustness: Wet-bulb Temperature and Primary Sectoral Employment



(a) Fourth-order polynomials



(b) Cumulative temperature bins



(c) Degree day bins

NOTES: This figure shows that the relationship between wet-bulb temperature and primary sectoral employment share is robust to alternative functional forms of temperatures. Each graph represents a predicted sectoral employment share-temperature response function (equation 1). Shaded areas are 95% confidence interval where robust standard errors are clustered at the province level.

likely than older workers to take up a job in formal non-agriculture. On the other hand, informal non-agriculture plays equally important role in absorbing workers of all groups. Again, there is no labor supply effect for any group, which suggests that this is not an income channel effect, else labor supply should rise.

Table 2: Wet-bulb Temperature and Primary Sectoral Employment: Heterogeneity by Age Group

Panel Approach

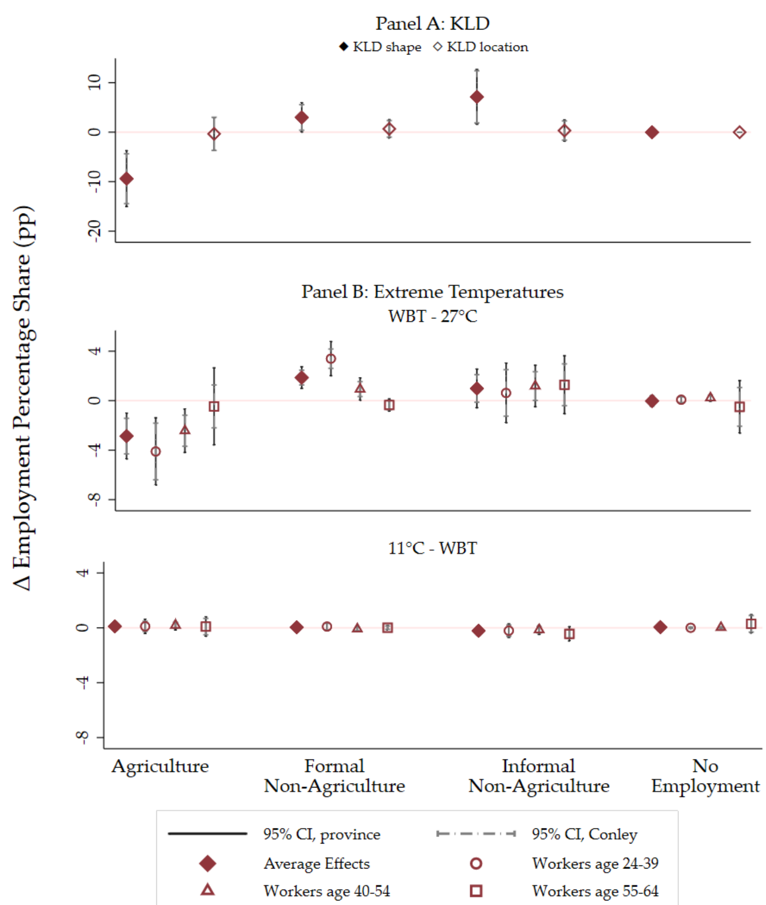
	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture	No Employment
	(1)	(2)	(3)	(4)
(WBT − 27) × Age 24-39 (G1)	-0.816	0.424	0.455	-0.060
	(0.226)	(0.150)	(0.172)	(0.072)
.	[0.229]	[0.165]	[0.161]	[0.061]
(WBT − 27) × Age 40-54 (G2)	-0.541	0.217	0.332	-0.021
	(0.119)	(0.107)	(0.154)	(0.058)
.	[0.202]	[0.079]	[0.165]	[0.041]
(WBT − 27) × Age 55-64 (G3)	-0.019	-0.063	0.234	-0.143
	(0.290)	(0.074)	(0.167)	(0.298)
.	[0.234]	[0.056]	[0.173]	[0.208]
p-value (G1) = (G2)	0.161	0.117	0.278	0.639
p-value (G2) = (G3)	0.038	0.002	0.570	0.679
p-value (G1) = (G3)	0.002	0.001	0.278	0.741
p-value (G1) = (G2) = (G3)	0.009	0.001	0.460	0.885
Observations	1707	1707	1707	1707
Province × Age Group FE	x	x	x	x
Region × Age Group × Year FE	x	x	x	x

NOTES: Unit of analysis is province-age group-year. Dependent variables are percentage shares of employment in each sector. All columns control for the second-order polynomials of precipitation and number of days with high wind speeds during the 12-month exposure interacted with age group dummies. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. Test of significant age cohort differences use standard errors clustered at the province level, results are qualitatively similar when using Conley standard errors. All regressions use sampling weights.

4.2.2 Long Differences Approach

Figure 4 presents the temperature effects on the four key outcomes estimated using the long differences approach. Controlling for extreme weather events and region-specific age group-specific trends, provinces experiencing a larger change in the shape of the temperature distribution between 1992-2006 and 2007-2018 see

Figure 4: Wet-bulb Temperature and Sectoral Employment: Long Differences Approach



NOTES: This figure presents the effects of temperature change, proxied by KLD shape and location difference (Panel A, evaluated at the sample mean values of temperature variables) and extreme temperatures (Panel B), on sectoral employment shares, which are obtained from estimating equation (3). Unit of analysis is province-age group. All regressions control for episodes of extreme precipitation (second-order polynomials in Panel B) and days with high wind speeds. Province distances are computed from province geographic centroids.

a larger reduction in the agricultural labor share and increases in both formal and informal non-agricultural employment shares, and these effects are all statistically significant at the 5% level. No temperature effect on the share of inactive and unemployed workers is detected (Panel A).

By using the shape and location components in the long differences specification, one can also examine the relative importance of "general warming" versus "increased risk of extreme temperatures" in inducing labor allocation. As seen, while the effects of shape difference are significant, general warming as proxied by location shifts virtually does not affect sectoral employment shares. The point estimates of location effects across specifications are small and close to zero. These findings are consistent with previous ecology literature, which emphasizes the potentially greater risks associated with changes in variation, as opposed to temperature mean, to the ecological systems (Vasseur et al. 2014; Turner et al. 2020).

In Panel B, I use changes in extreme temperatures during the reference period as the key independent variables of interest. The effects therefore can be directly compared with those estimated with equation (1). Consistent with findings from panel approach, cold temperatures generally do not have any significant effect on sectoral employment shares. Across agriculture and formal non-agriculture specifications, the long differences estimates of hot temperature effects are of larger magnitude than the respective panel estimates. Evaluated at the sample mean, the point estimate of the hot temperature effects on agricultural labor share is -2.80 (95% CI = [-4.36, -1.24]) with the panel approach, and -4.65 [-7.68, -1.62] with the long differences approach. The corresponding effects on formal non-agricultural labor share are 1.28 [0.29, 2.27] (panel approach), and 2.85 [1.42, 4.28] (long differences approach). There is no difference between the effect of hot temperatures on informal non-agricultural labor share across approaches, however.[34]

Similar to the panel approach, the long differences estimation also yields significant heterogeneous effects of hot temperatures on agricultural and formal non-agricultural labor shares across age groups. There is no evidence of differential temperature effects across age groups for informal non-agriculture, as well as no labor supply effect.[35]

4.3 Robustness Checks, Placebo Test, and Additional Analyses

Robustness Checks I perform a series of robustness checks in Appendix B. To summarize, the results are robust to different sample restrictions, different specifi-

[34]The panel and long differences estimates for informal non-agriculture are 1.78 [0.45, 3.12] and 1.80 [-0.63, 4.23], respectively.

[35]The F-test for significant age cohort differences yields p-values of 0.12, 0.00, 0.68, and 0.42 for the specification of agriculture, formal non-agriculture, informal non-agriculture, and no employment, respectively.

cations (with the inclusion of time-varying demographic characteristics including educational attainment, share of ethnic minority population, share of male population, as well as to the inclusion of lagged dependent variables), to a different method of constructing weather variables (averaging values of grid points over a geographical boundary), to different weather exposures. Results from the long differences approach are generally robust to different period definitions. I also present additional results using dry-bulb temperatures, which are qualitatively similar to those obtained from the main analysis.

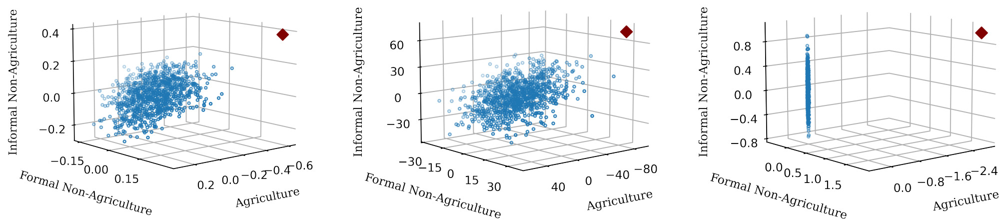
Placebo Test As an additional check on the econometric specifications, I conduct a placebo test with Monte Carlo analyses of equations (1) and (3) using actual employment and climate data to ensure that the panel and long differences approaches provide correct inference and unbiased estimates. Specifically, in each Monte Carlo iteration, I randomly reassign the weather series from one province-age group unit to another province-age group's employment series, and then test for temperature effects in equations (1) and (3). The idea is that incorrect assignment of weather distribution to province-age group employment shares should yield results of smaller magnitude with zero mean or different sign and statistical insignificance. Figure 5 presents the joint distribution of the estimated coefficients with random assignment for agricultural, formal, and informal non-agricultural employment share outcomes. As seen, the set of baseline estimates fall far outside the resulting joint distribution of spurious random reassignment estimates, suggesting that the temperature-sectoral employment share relationship is unlikely to arise by chance. The observed Type I error rates across all approaches-sectoral outcomes are approximately 4-6% when evaluating at the 5% significance level. These findings suggest that the inference is fairly accurate against the null hypothesis of no temperature effect.

Migration Analysis In studying intersectoral labor reallocation, I have implicitly assumed that local markets are bounded at the province level. Previous works, however, have demonstrated the prevalence of human migration across spaces in response to climate change.[36] Inter-provincial migration might alter demographic compositions and therefore mechanically lead to changes in sectoral employment shares at the province level, without being driven by underlying forces as will be

[36]See Cattaneo et al. (2019) for a review of relevant literature.

Figure 5: Placebo Test: Wet-bulb Temperature and Primary Sectoral Employment

Results from Monte Carlo Permutation Tests



(a) Panel Approach (b) Long Differences: KLD (c) Long Differences: WBT

NOTES: Each graph presents the (joint) distribution of estimated coefficients (blue circles) of temperature effects on sectoral employment shares, which are obtained from 1,000 Monte Carlo simulations where the weather series of one province-age group unit is randomly reassigned to another province-age group's employment share series. Red diamonds represent the baseline estimates from the panel approach (Panel a), the long differences approach using KLD measures (Panel b), and the long differences approach using extreme temperature measures (Panel c).

shown in Section 5. While I do not have migration data for each age group over the study period, in Appendix Table C1, I provide supporting evidence showing that both cold and hot temperatures do not significantly affect the rates of migration, including out-migration, in-migration, and net-migration at the province level, thereby suggesting that the provincial-level labor markets are relatively bounded, and within-province intersectoral labor reallocation is an empirically relevant margin in this setting.

Gender and Education Analyses Previous literature has demonstrated the potentially differential effects of environmental changes in general on human capital and labor outcomes by gender (Maccini and Yang 2009; Björkman-Nyqvist 2013). In Appendix Table C5, I find no evidence of heterogeneous effects of hot temperatures on intersectoral labor reallocation by gender. Likewise, I find limited evidence of differential temperature effects by education level, which suggests that the heterogeneous results by age cohorts in the formal non-agricultural sector cannot be entirely explained by differences in educational attainment across these groups.

5 Potential Mechanisms

The analysis so far yields two main results. First, temperature changes, particularly at the higher end of the distribution, accelerate a movement of workers out of agriculture. Second, there are heterogeneous temperature effects across age groups and sectors of destination work. This section explores potential mechanisms underlying these results.

5.1 Mechanism: Hot Temperatures Accelerate Labor Reallocation

The results on average effects where hot temperatures induce reallocation of workers from agriculture to non-agricultural sectors are consistent with being dominantly driven by the relative labor productivity loss mechanism. In what follows, I offer additional evidence to support this channel.

Reallocation effects are more pronounced in areas that are more integrated into the world market and whose prices are less affected by temperature change. To begin, I test whether hot temperatures affect labor allocation in areas with decent access to trade more than in distant areas. I estimate a variant of equations (1) and (3) where weather variables, including temperature, are interacted with a "tradable" indicator. I proxy for tradability using two measures. The first measure is the distance from a province geographic centroid to the nearest major seaport.[37] The second measure is the correlation coefficient between local agricultural price series, specifically rice price, and that of the world market.[38] The idea is that areas closer to major seaports and/or more integrated to the world market are less affected by temperature-induced price and local demand effects, and thus most labor reallocation is driven by relative labor productivity loss.

Table 3 shows that the effects on agricultural and formal non-agricultural employment shares are driven by relatively tradable areas. In provinces that are less connected, there were actually opposite effects: hot temperatures are associated with an increase in agricultural labor share and decreases in non-agricultural labor shares. Even though these effects are imprecisely estimated, the point estimates of the temperature effects on formal and informal non-agricultural sectors further suggest that local demand effects play an important role in these isolated areas:

[37] The three major seaports considered include Hai Phong, Da Nang, and Sai Gon.

[38] For details in the construction of this measure, see Appendix A3.

Table 3: Wet-bulb Temperature and Primary Sectoral Employment across Space

	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture	No Employment
	(1)	(2)	(3)	(4)
Panel A: Tradability proxied by distance to the nearest major port				
(WBT - 27) × (Tradable = 0) (N)	3.816	-0.181	-3.956	0.401
	(2.802)	(1.338)	(2.171)	(0.727)
.	[5.134]	[1.637]	[2.622]	[2.572]
(WBT - 27) × (Tradable = 1) (T)	-0.602	0.282	0.380	-0.061
	(0.167)	(0.104)	(0.143)	(0.083)
.	[0.198]	[0.099]	[0.161]	[0.065]
p-value (N) = (T)	0.122	0.731	0.052	0.520
Panel B: Tradability proxied by correlation coefficient between local and world market price of rice				
(WBT - 27) × (Tradable = 0) (N)	-0.095	0.240	-0.178	0.027
	(0.222)	(0.131)	(0.152)	(0.075)
.	[0.248]	[0.120]	[0.224]	[0.069]
(WBT - 27) × (Tradable = 1) (T)	-0.558	0.243	0.373	-0.058
	(0.185)	(0.107)	(0.158)	(0.079)
.	[0.179]	[0.105]	[0.136]	[0.062]
p-value (N) = (T)	0.013	0.977	0.000	0.076
Observations	1707	1707	1707	1707
Province × Age Group FE	x	x	x	x
Region × Age Group × Year FE	x	x	x	x

NOTES: Unit of analysis is province-agegroup-year for the panel approach and province-agegroup for the long differences approach. Dependent variables are percentage shares of employment in each sector. In Panel A, "Tradable" is an indicator that takes value 1 if the distance from a province centroid to the nearest major port is below the 75th percentile (approximately 240 km) and 0 otherwise. In Panel B, "Tradable" is an indicator that takes value 1 if the correlation coefficient between local median rice price and world rice price series is above the 25th percentile (approximately 0.67) and 0 otherwise. All regressions control for weather variables (second-order polynomials of precipitation and wind speed) and their interactions with 'Tradable' dummy. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

there is a much larger decrease in share of workers in informal non-agriculture, whose products are mostly non-tradable. These findings are all consistent with the finding of no labor supply effect at the extensive margin.

Hot temperatures have disproportionately negative effect on agricultural labor supply. Second, I test whether hot temperatures affect labor supply, as proxied by the number of hours worked. I estimate model (1) where the dependent variable is the number of hours worked in each sector for the sample of individuals working in that sector. If hot temperatures affect human health and task performance, it might increase labor dis-utility and lead to a reduction in their labor supply (Rode et al. 2022).

Table 4 presents the effects of hot temperatures on average yearly hours of work estimated from the panel model. In Panel A, the dependent variable is the mean of hours of work, conditional on working in a sector. In Panel B, the dependent variable is mean of hours worked of individuals in an analysis unit, regardless of whether an individual workers in a specific sector or not (i.e., individuals not working in a specific sector is considered working zero hours). The results in Table 4 imply that one extra degree day higher than 27°C wet-bulb temperature decreases both measures of hours worked in agriculture by approximately 20-30 hours per year. Given that the average cumulative exposure higher than the 27°C threshold is 4.7 degree days per year, an average agricultural worker tends to lower their labor supply by 98-140 hours per year, or 20-30 minutes per day worked (intensive and extensive margin effects). Hours of work of existing workers in formal and informal non-agricultural sectors, however, are generally not affected by hot temperatures. These results imply that the increase in unconditional hours of work in formal non-agriculture (Columns 2 and 3, Panel B) is largely driven by new workers switching to this sector in response to hot temperatures over time (extensive margin effects). There is some suggestive evidence of a decline in total labor supply (Column 4, Panel B), although the effect is insignificant at conventional levels.

Hot temperatures have disproportionately negatively affected agricultural labor productivity. The results on conditional hours of work suggest that hot temperatures cause a reduction in labor inputs to agricultural production but not formal or informal non-agricultural production, which is consistent with findings from Graff Zivin and Neidell (2014), who show that temperature increases at the right tail of the distribution reduce hours worked in climate-highly exposed industries. If such a response translates into sectoral productivity loss, hot temperatures can have differential effects on relative labor productivity loss across sectors.

I directly test the heterogeneous temperature effects on sectoral labor produc-

Table 4: Hot Wet-bulb Temperature and Labor Supply

Hours of work are computed for the primary and secondary jobs

	Panel A: Conditional Hours of Work (Intensive Margin)		
	(1) Agriculture	(2) Formal Non-Agriculture	(3) Informal Non-Agriculture
(WBT – 27)	-30.433	2.747	0.193
	(8.608)	(4.820)	(4.280)
.	[9.744]	[3.804]	[4.106]
Sample Mean of DepVar	1297	1884	1768

	Panel B: Unconditional Hours of Work (Extensive Margin)			
	(1) Agriculture	(2) Formal Non-Agriculture	(3) Informal Non-Agriculture	(4) Total
(WBT – 27)	-20.972	6.362	5.815	-8.893
	(4.599)	(2.481)	(2.872)	(5.791)
.	[5.579]	[2.446]	[3.252]	[6.081]
Sample Mean of DepVar	766	405	667	1846
Observations	1551	1551	1551	1551
Province × Age Group FE	x	x	x	x
Region × Age Group × Year FE	x	x	x	x

NOTES: Unit of analysis is province-age group-year. In 2002, only hours worked for the primary job are recorded and thus data from VHLSS 2002 are dropped for consistency. Dependent variables are average number of hours worked, winsorized at the top 1% of the individual distribution by year. All columns control for the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. All regressions use sampling weights.

tivity. Specifically, I assemble a longitudinal dataset of province-level production for agriculture, formal non-agriculture, and informal non-agriculture and estimate the effect of hot temperatures on revenue per worker of each sector.[39] Table 5 shows

[39]Ideally, one should estimate the marginal product of labor in each sector and then examine the effect of hot temperatures on that outcome. Details on such an approach are available in Appendix D. Data limitations, however, do not allow me to estimate the marginal product of labor.

I estimate the effect of hot temperatures on revenue per worker with a year-to-year panel regression that the effect of hot temperatures is significantly larger in magnitude for labor productivity in agriculture than other sectors of the economy. In particular, one extra degree day higher than 27°C leads to a 1.2% decrease in revenue per worker in agriculture (p-value < 0.01). The corresponding effects on formal and informal non-agricultural labor productivity are close to zero and statistically insignificant at conventional levels. Appendix Table C3 further shows that hot temperatures also negatively affect agricultural yields but the effects are smaller in magnitude relative to labor productivity: one degree higher than 27°C leads to approximately 0.4-0.7% decrease in yields of rice–the main staple crop. This finding suggests that heat's impact on agricultural labor transcend the commonly studied land productivity mechanism wherein lower crop yields drive labor reallocation out of agriculture.

Although hot temperatures do not affect formal non-agricultural labor productivity on average, a subset of non-agricultural workers is also adversely affected. Appendix Table C4 presents additional results using an unbalanced 15-year longitudinal dataset of firms that appeared at least twice during the period 2002-2016. The effects of hot temperatures on labor productivity in climate highly exposed industries such as mining and quarrying are as large as in agriculture. For example, one degree day above 27°C is associated with approximately 1.6% decrease in labor productivity of small and old mining firms. This further supports the underlying mechanism being a reduction in human labor productivity when workers are exposed to thermal stress.

Taken together, these findings suggest that the relative labor productivity loss mechanism dominates and year-to-year variation in hot temperatures induce workers to move out of agriculture. The fact that we observe similar results in the effects of hot temperatures on labor reallocation both in the short run and in the long run

---

of the following form:

$$y_{prt} = f(\mathrm{WBT}_{pt}) + g(\mathrm{R}_{pt}) + \gamma_p + \gamma_{rt} + \varepsilon_{prt}$$

where y_{prt} denotes the log of revenue per worker in each sector (agriculture, informal non-agriculture, and formal non-agriculture) in province p of region r in year t. The term R_{pt} represents a vector of other weather variables in province p in the reference period relative to year t, including second-degree polynomials of rainfall and episodes of high speed wind. The vector γ_p represents province-specific fixed effects, which control for province-specific time-invariant unobserved characteristics that can affect the outcome. The term γ_{rt} denotes region-specific year fixed effects, which is to control for aggregate-level shock at the economic region level that is time-varying. Again, I cluster the standard errors at the province level and also report Conley standard errors that allow for spatial correlation up to 200 km and temporal correlation up to five lags. Details on the construction of this analysis dataset can be found in Appendix A1.1.

Table 5: Wet-bulb Temperature and Sectoral Labor Productivity: 2002-2016

	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture
	(1)	(2)	(3)
(WBT – 27)	-0.011985	0.006646	-0.001826
	(0.003504)	(0.008571)	(0.006487)
.	[0.002636]	[0.008169]	[0.010917]
Sample Mean of DepVar	2.28	5.29	3.51
Province FE	x	x	x
Region-by-Year FE	x	x	x
Observation	416	3364	416

NOTES: Unit of analysis is province-year. Dependent variables are log of annual revenue per worker (2010 million VND). All columns control for cold temperatures (cumulative wet-bulb temperature less than 11°C), the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. All regressions use production size (number of workers) as weights.

in Section 4.2 suggests that the labor productivity mechanism likely holds in the long run as well. This finding is consistent with classic predictions of small open economy models, where changes in relative labor productivity loss and thus comparative advantage induce workers to move away from the less productive sector to relatively more productive sectors. It contrasts with Liu, Shamdasani, and Taraz (forthcoming), which shows an intensification of local demand effects over a longer time frame and thus rising temperatures end up leading to a reduction in non-agricultural labor shares over longer time scales in Indian districts.

5.2 Mechanism: Differential Effects by Age Group

The second set of main results is that hot temperatures have differential effects on the rate of reallocation by age group and sector into which workers move. In this section, I use a simple Roy-Borjas framework to show how the differential temperature effects on labor reallocation by age group can be explained by the existence of non-uniform labor market frictions across the two dimensions: sector of employment and age. In particular, when hot temperatures cause changes in relative labor productivity and thus return to labor, workers are induced to move to sectors that entail lower switching costs. I then use a panel dataset of individual workers from different age groups, who move across sectors over years, to infer the extent of frictions, showing that the data are consistent with this framework.

The Roy-Borjas Framework This model provides a useful framework to infer frictions from gaps in sectoral outcomes using individual-level panel data (Roy 1951; Borjas 1987). Because the share of no employment (including both unemployed and inactive) remains relatively stable over the study period, and is not significantly affected by temperature changes, for simplicity, let us assume full employment in each of the two sectors $s \in \{g, n\}$, where g represents agriculture and n denotes either formal or informal non-agriculture. Let $\ln w_{js}$ be the (log) earnings of individual j in sector s, which can be decomposed into two parts: the part explained by sector-specific returns to observable characteristics μ_s, and the part due to unobserved characteristics ϵ_{js}.

$$\ln w_{js} = \mu_s(\text{WBT}) + \epsilon_{js} \tag{5}$$

where $\epsilon_{jg} \sim N(0, \sigma_g^2)$, $\epsilon_{jn} \sim N(0, \sigma_n^2)$, and are assumed to be jointly normally distributed with covariance $\text{cov}(\epsilon_{jg}, \epsilon_{jn}) = \sigma_{g,n}$. One can think of μ_s as workers in the same sector who share similar observable characteristics and receive an average wage rate based on such characteristics. Any difference in the average wages across sector can be driven by selection on and differential returns to observables. For example, previous literature has shown that growing sectors during structural change are more skill-intensive and have higher returns to education (Herrendorf and Schoellman 2018; Buera, Kaboski, and Zhao 2019). Similarly, female workers may have comparative advantage in service and some manufacturing industries (Ngai and Petrongolo 2017). The differential returns to individual characteristics at the firm-level can be considered the result of firm-specific distortions, such as scale effects that impact resource reallocation among firms within a sector (Adamopoulos and Restuccia 2014; Donovan 2021). In this paper, I distinguish these from barriers that prevent the free flow of workers across sectors, which is to be explained next.

The sector-specific average return μ_s is assumed to depend on the temperature distribution, and more specifically, hot temperatures. As shown in Section 5.1, the labor productivity of agricultural workers is disproportionately affected by hot temperatures relative to non-agricultural workers, which can be expressed mathematically as:

$$\frac{\partial \mu_g(\text{WBT})}{\partial \text{WBT}} < \frac{\partial \mu_n(\text{WBT})}{\partial \text{WBT}} \leq 0 \tag{6}$$

The following arguments are at the individual level so subscript j is dropped. Assume each worker starts in agriculture, and let C be the cost of labor market frictions associated with switching sector of employment. Specifically, non-agricultural workers effectively receive a fraction of their earnings $w_n - C = w_n(1-\tau)$ where $\tau_n \in [0,1)$ denotes the extent of frictions. Some examples of frictions that can be accounted for by τ include (i) taxes and social security contribution that are specific to (formal) non-agriculture, (ii) psycho-social costs associated with "transferring from easy going way of life of subsistence to more regimented environment" (Lewis 1954), (iii) cost of acquiring information on job opportunities, (iv) cost associated with sector-specific skill or educational investment, as well as (v) mobility cost.[40] In other words, C captures not only one-off search or switching costs, but also the discounted net present value of costs arising due to recurring frictions.

The worker's decision to switch out of agriculture is determined by the sign of the index function:

$$\mathbb{I} = \ln\left(\frac{w_n(1-\tau)}{w_g}\right) \approx \mu_n(\text{WBT}) - \mu_g(\text{WBT}) - \tau + \epsilon_n - \epsilon_g \tag{7}$$

such that the individual switches if $\mathbb{I} > 0$ and does not switch if $\mathbb{I} < 0$. Let $\nu = \epsilon_n - \epsilon_g$ and $\Phi(.)$ be the CDF of the standard normal then the switching rate is:

$$\Pr\left\{\nu > -\left[\mu_n(\text{WBT}) - \mu_g(\text{WBT}) - \tau\right]\right\} = 1 - \Phi\left[\frac{\mu_g(\text{WBT}) - \mu_n(\text{WBT}) + \tau}{\sigma_\nu}\right] \tag{8}$$

which is increasing in relative returns to observed characteristics, and decreasing in costs associated with frictions. This expression, combined with equation (6), implies that all else constant, when there is an increase in hot temperatures and thus change in relative returns to labor, workers are more likely to reallocate from agriculture to non-agricultural sectors, which is the first main finding of this paper.

A general interpretation of equation (8) is that changes in relative returns to ob-

[40] Mobility cost occurs because non-agricultural jobs are often concentrated in urban areas whereas agricultural jobs are in rural areas, although this might not be necessarily true in the case of informal non-agricultural jobs.

served characteristics induce workers to move towards sectors or occupations that entail lower switching costs. While some of the above-mentioned sector-specific frictions such as labor taxes are uniform across individuals, others such as skills investment and re-training for a formal job might not be. For older workers, investment in acquiring a new set of skills may be relatively more costly than for younger workers due to differential opportunity costs of time and work-life horizons. Thus, all else constant, such frictions imply lower switching rates out of agriculture for older workers relative to younger workers.

To demonstrate the role of frictions in the observed sectoral earnings gaps, let us compare the average earnings in agriculture and non-agriculture of individuals who switch from agriculture to non-agriculture, which is the dominant switching direction for individuals who do switch (Hamory et al. 2021; Herrendorf and Schoellman 2018). In practice, we do not observe $\mathbb{E}\left(\ln w_g | \mathbb{I} > 0\right)$. If we are to approximate it with the earnings of agricultural workers right before their transition to non-agriculture, then the observed sectoral gap in earnings among switchers would be:

$$\begin{aligned} \text{gap} &\approx \mathbb{E}\left(\ln w_n | \mathbb{I} > 0\right) - \mathbb{E}\left(\ln w_g | \mathbb{I} > 0\right) \\ &= \underbrace{\mu_n - \mu_g}_{\substack{\text{different returns}\\ \text{to observables}}} + \underbrace{\left[\frac{(\sigma_n - \sigma_g)^2 + 2\sigma_n\sigma_g(1-\rho_{g,n})}{\sigma_\nu}\right] \frac{\phi\left(\frac{\mu_g - \mu_n + \tau}{\sigma_\nu}\right)}{1 - \Phi\left(\frac{\mu_g - \mu_n + \tau}{\sigma_\nu}\right)}}_{\substack{\text{selection on unobservables}\\ \text{labor market frictions}}} \end{aligned} \tag{9}$$

where $\phi(.)$ is the PDF of the standard normal distribution, and $\rho_{g,n}$ is the correlation coefficient of the productive ability in agriculture ϵ_g with productive ability in non-agriculture ϵ_n. Equation (9) illustrates that the observed earnings gap depends on (i) differential returns to observed individual characteristics (such as education) across sectors, (ii) differences in unobservables (selection on unobservables), and (iii) labor market frictions.[41]

The following arguments are in general equilibrium. Consider the first case when there is no selection and no friction, all workers have the same productive ability across sectors: $\epsilon_g = \epsilon_n$ and $\rho_{g,n} = 1$, a necessary condition for an interior solution to sectoral labor allocation is $\mu_n = \mu_g$, that is, there is no differences in

[41] With the sample of switchers, selection on observables is minimized. The first term of equation (9) reflects differential returns to observables.

returns to individual characteristics across sectors. As a result, $\mathbb{E}(\ln w_n|\mathbb{I} > 0) - \mathbb{E}(\ln w_g|\mathbb{I} > 0) = 0$. Similarly, when there are frictions and no selection, a necessary condition for an interior solution to sectoral labor allocation is $\mu_n - \mu_g = \tau$. The earnings gap is non-zero and equal to the costs associated with frictions: $\mathbb{E}(\ln w_n|\mathbb{I} > 0) - \mathbb{E}(\ln w_g|\mathbb{I} > 0) = \tau$. When there are selections and no friction, the second term of equation (9) is strictly positive. In this case, the earnings gap depends on equilibrium prices/returns to observed individual characteristics across sectors, how dispersed the distributions of unobservables are, and how strong the values of unobserved skills are correlated across sectors.

These three cases illustrate that the existence of sectoral earnings gap among the sample of switchers is consistent with both frictions and selection. Under the assumption that workers who switch are nearly indifferent between the two sectors and are induced to switch at the margin because of change in relative returns to observables (Schoellman 2020), then the gains in earnings reflect the extent of frictions.[42]

Taken together, the framework and the assumption of marginal switchers imply the following. First, if the gains from moving from agriculture to informal non-agriculture are insignificant and do not differ across age groups, then it is evident that there is little cost of switching from agriculture to informal non-agriculture and thus temperature-driven change in relative return to labor will induce workers of all age groups to move into informal non-agriculture at a similar rate. Second, if the gains from moving from agriculture to formal non-agriculture are significant and differ across age groups, then it is suggestive that there is a large cost of switching from agriculture to formal non-agriculture, where the older the workers, the larger the cost they will incur. As a result, only younger workers, who incur less switching costs, will take up a job in the formal non-agricultural sector in response to hot temperatures.

Indirect Evidence of Non-Uniform Labor Market Frictions Guided by the above framework, I infer the extent of frictions by estimating the following regression

[42] Note that in this framework, without this assumption, the size of the gains in earnings is generally not informative about the extent of frictions because an increase in τ can result in ambiguous effects on earnings gaps, depending on the relationship between frictions and relative returns to observables. When $\tau > \mu_n - \mu_g$, an increase in τ results in decreases in both $\phi(.)$ and $1 - \Phi(.)$ and thus the net effect is ambiguous. When $\tau < \mu_n - \mu_g$, an increase in τ leads to an increase in $\phi(.)$ and a decrease in $1 - \Phi(.)$, resulting in a larger gap.

using the sample of sector-switchers from a pool of three-consecutive-survey-wave individual panel data sets over the period 2002-2018:

$$\ln w_{jt} = \sum_{a} \sum_{s} \alpha_{a,s} \mathbb{I}_a \mathbb{I}_s + \psi Z_{jt} + \gamma_j + \gamma_t + \varepsilon_{jt} \tag{10}$$

where $\ln w_{jt}$ is the log real earnings of individual j at time t. Earnings are computed as the sum of labor wages, benefits, as well as household farm or non-farm net profits.[43] The term $\mathbb{I}_s$ takes the value of one if the individual works in sector s for their main job and zero otherwise. There are three sectors: agriculture (g), formal non-agriculture (f) and informal non-agriculture (i). The term $\mathbb{I}_a$ denotes whether the individual j belongs to age group $a \in \{24-39, 40-54, 55-64\}$ that corresponds to the three age groups in the main empirical analysis. Other time-variant controls such as age, age squared, and log hours worked are included in vector Z. The vectors γ_j and γ_t denote individual and year fixed effects, respectively. Individual fixed effects are important because they control for individual-specific time-invariant unobservables as well as time-invariant observables such as gender, ethnicity, or educational attainment, thereby minimizing the role of self-selection. Under the assumption that switchers are marginal workers, the difference between the parameters $\alpha_{a,s\in\{f,i\}}$ and $\alpha_{a,g}$ reflect the extent of frictions in formal and informal non-agriculture for switchers from agriculture for each age group. I also estimate a variant of equation (10), where individual-specific time-invariant characteristics, including education level, gender, and ethnic minority indicator, are interacted with sector dummies, thereby allowing the returns to time-invariant observables to vary across sectors.[44]

Table 6 presents the results from estimating equation (10). First, there are large gains in annual average earnings for workers who switched from agriculture to informal and formal non-agriculture, even after controlling for hours worked and individual-specific time-invariant unobserved characteristics. Under the assumption that the returns to individual observables are uniform across sectors, workers switching from agriculture to informal non-agriculture earn 20% more on average, and the gain is 30% if they transition into formal non-agriculture (Column 1). When

[43] Household members are assumed to receive a share of household net profits that is proportional to their hours worked in household farm and non-farm business.

[44] I do not interact time-varying characteristics with sector dummies to reflect the idea that it is purely the change in relative returns to observables, not any change in individual characteristics or their behaviors, that induces workers to switch sectors.

Table 6: Gains in Labor Earnings among Switchers

	Log(Earnings) inc. profits	
	(1)	(2)
gains in moving from agriculture to informal non-agriculture		
24-39 (G1I)	0.249	0.161
	(0.015)	(0.132)
40-54 (G2I)	0.223	0.144
	(0.015)	(0.132)
55-64 (G3I)	0.274	0.212
	(0.035)	(0.136)
gains in moving from agriculture to formal non-agriculture		
24-39 (G1F)	0.335	0.525
	(0.018)	(0.126)
40-54 (G2F)	0.307	0.521
	(0.019)	(0.126)
55-64 (G3F)	0.456	0.673
	(0.048)	(0.133)
p-value G1I = G2I	0.201	0.423
p-value G2I = G3I	0.171	0.068
p-value G1I = G3I	0.507	0.176
p-value G1F = G2F	0.266	0.883
p-value G2F = G3F	0.003	0.003
p-value G1F = G3F	0.018	0.004
p-value G1I = G2I = G3I	0.253	0.179
p-value G1F = G2F = G3F	0.013	0.009
adj. R^2	0.545	0.550
Observations	37642	37642
Individuals	14040	14040
Year FE	x	x
Individual FE	x	x
Individual Controls $\times$ Sector		x

NOTES: Sample includes workers who switched sector at least once in each three-wave panel. All regressions control for log hours worked, age and age squared. Earnings include labor wages, other benefits and household farm/non-farm profits, trimmed at its top and bottom 5%. Individual controls include gender, ethnicity, marital status (married, single, widowed/separated), and general education qualification (no education, primary education, lower secondary education, upper secondary education, post secondary education). Robust standard errors clustered at individual level are in parentheses. SOURCES: Data from VHLSS three-wave individual panel datasets 2002-2004-2006, 2004-2006-2008, 2010-2012-2014, 2012-2014-2016, and 2014-2016-2018.

the uniform return assumption is relaxed, the residual gains when moving to informal non-agriculture are less precisely estimated, whereas the gains from moving to formal non-agriculture remain similar in magnitude and statistical significance (Column 2).[45] If switchers are thought of as being marginal workers and thus there is no selection bias involved, then the above framework suggests that the costs associated with frictions in informal non-agriculture are smaller compared to formal non-agriculture.

Second, across the two specifications, there is no differential gains across age groups when moving from agriculture to informal non-agriculture, which, according to the Roy-Borjas framework, suggests the existence of a uniform switching cost across age groups if they transition into this sector. On the other hand, the group of older workers (age 55-64) experience the largest gains when switching into formal non-agriculture, whereas there is little evidence of differential gains between the younger two groups. Viewed through the lens of the above framework, this evidence is suggestive that the older the workers, the larger the switching costs they incur, which makes them less likely to move out of agriculture into formal non-agriculture in responses to relative labor productivity loss caused by hot temperatures.

6 Conclusions

Climate change and associated extreme weather events affect different aspects of the economy. Earlier works show that under negative agricultural productivity growth induced by weather shocks and temperature rises, we observe reallocation of workers away from and into agriculture. In this paper, I add to the discussion by showing that more than the commonly-studied land or crop productivity mech-

[45]Under both assumptions, there is no full first-order dominance between sector-specific residual gains among workers. However, a smaller (larger)percentage of workers have negative (positive) gains if transitioning into both informal and formal non-agriculture relative to agriculture (See Appendix Figure C5). I also find roughly 36-52 log-point differences in earnings from cross-sectional analysis under the assumption of uniform return to individual observables across sectors (with agriculture-formal non-agriculture gaps being the largest) after adjusting for individual controls and hours worked. In general, even though these cross-sectional estimates are smaller than those reported by Gollin, Lagakos, and Waugh (2014), they are 40-60% larger than those obtained from sample of switchers with individual fixed effects reported in Column (1) of Table 6. These observations are not specific to this setting but found in other settings as well. Hamory et al. (2021)'s work on Indonesia and Kenya, and Alvarez (2020)'s work on Brazil are among several papers on low and middle income countries.

anism could drive an outflow of labor from agriculture in the short and long runs. Specifically, workers are induced to move into sectors where their labor productivity is less affected by hot temperatures.

Although climate change accelerating the reallocation of labor away from the relatively low-productivity agricultural sector may sound beneficial to the economy, the fact that climate change also affects labor productivity in other sectors, and that a large part of such reallocation is into the informal non-agricultural sector makes the overall effects of climate change-induced labor reallocation less certain. As summarized by Ulyssea (2020), there are at least three complementary views regarding the role of informality: (i) informality as a reservoir of potentially productive entrepreneurs who are constrained by formal entry costs and regulations, (ii) informality as parasite firms that can survive in the formal sector but choose to remain informal to earn higher profits by avoiding taxes and regulations, and (iii) informality as a survival strategy for low-skill workers. If a large part of informal workers in the Vietnamese economy fall into the latter, survival category as in Brazil (Ulyssea 2018), then such a reallocation might have important welfare consequences by reinforcing the country's comparative advantage in less skill-intensive industries, which, if combined with low rates of innovation, might lead to lower long run growth (Bustos et al. 2020).

Finally, hot temperatures have differential effects on labor reallocation across age groups and sectors into which workers move. This also has implications for the welfare distribution and within-country inequality consequences of climate change. A promising avenue for future research is to evaluate the welfare consequences of climate change-induced labor reallocation on different age groups, and of the economy as a whole.

References

Adamopoulos, T. and D. Restuccia (2014). "The size distribution of farms and international productivity differences". *American Economic Review* 104.6, pp. 1667–1697.

Albert, C., P. Bustos, and J. Ponticelli (2021). *The Effects of Climate Change on Labor and Capital Reallocation*. NBER Working Paper w28995. National Bureau of Economic Research.

Alvarez, J. A. (2020). "The Agricultural Wage Gap: Evidence from Brazilian Microdata". *American Economic Journal: Macroeconomics* 12.1, pp. 153–173.

Amari, S.-i. (2016). *Information geometry and its applications*. Vol. 194. Springer.

Auffhammer, M., S. M. Hsiang, W. Schlenker, and A. Sobel (2013). "Using weather data and climate model output in economic analyses of climate change". *Review of Environmental Economics and Policy* 7.2, pp. 181–198.

Barrett, C. B., A. Ortiz-Bobea, and T. Pham (forthcoming). "Structural transformation, agriculture, climate and the environment". *Review of Environmental Economics and Policy*.

Bathiany, S., V. Dakos, M. Scheffer, and T. M. Lenton (2018). "Climate models predict increasing temperature variability in poor countries". *Science Advances* 4.5, eaar5809.

Beck, H. E., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg, and E. F. Wood (2018). "Present and future Köppen-Geiger climate classification maps at 1-km resolution". *Scientific Data* 5.1, pp. 1–12.

Behrer, A. P. and J. Park (2017). *Will we adapt? temperature, labor and adaptation to climate change*. Harvard Project on Climate Agreements Working Paper, pp. 16–81.

Björkman-Nyqvist, M. (2013). "Income shocks and gender gaps in education: Evidence from Uganda". *Journal of Development Economics* 105, pp. 237–253.

Bobba, M., L. Flabbi, S. Levy, and M. Tejada (2021). "Labor market search, informality, and on-the-job human capital accumulation". *Journal of Econometrics* 223.2, pp. 433–453.

Bonnet, F., J. Vanek, and M. Chen (2019). *Women and men in the informal economy: A statistical brief*. Report. WIEGO.

Borjas, G. (1987). "Self-Selection and the Earnings of Immigrants". *American Economic Review* 77.4, pp. 531–553.

Brauw, A. de and T. Harigaya (2007). "Seasonal migration and improving living standards in Vietnam". *American Journal of Agricultural Economics* 89.2, pp. 430–447.

Buera, F. J., J. P. Kaboski, and M. Q. Zhao (2019). "The rise of services: The role of skills, scale, and female labor supply". *Journal of Human Capital* 13.2, pp. 157–187.

Burke, M. and K. Emerick (2016). "Adaptation to climate change: Evidence from US agriculture". *American Economic Journal: Economic Policy* 8.3, pp. 106–40.

Bustos, P., B. Caprettini, and J. Ponticelli (2016). "Agricultural productivity and structural transformation: Evidence from Brazil". *American Economic Review* 106.6, pp. 1320–65.

Bustos, P., J. M. Castro-Vincenzi, J. Monras, and J. Ponticelli (2020). *Industrialization without Innovation*. NBER Working Paper w25871. National Bureau of Economic Research.

Carleton, T., A. Jina, M. Delgado, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, R. E. Kopp, K. E. McCusker, I. Nath, J. Rising, A. Rode, H. K. Seo, A. Viaene, J. Yuan, and A. T. Zhang (2022). "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits". *Quarterly Journal of Economics*. qjac020.

Cattaneo, C., M. Beine, C. J. Fröhlich, D. Kniveton, I. Martinez-Zarzoso, M. Mastrorillo, K. Millock, E. Piguet, and B. Schraven (2019). "Human migration in the era of climate change". *Review of Environmental Economics and Policy*.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019). *Inference with arbitrary clustering*. IZA Discussion paper 12854.

Colmer, J. (2021). "Temperature, labor reallocation, and industrial production: Evidence from India". *American Economic Journal: Applied Economics* 13.4, pp. 101–24.

Conley, T. G. (1999). "GMM estimation with cross sectional dependence". *Journal of Econometrics* 92.1, pp. 1–45.

Dell, M., B. F. Jones, and B. A. Olken (2014). "What do we learn from the weather? The new climate-economy literature". *Journal of Economic Literature* 52.3, pp. 2774–2814.

Donovan, K. (2021). "The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity". *Review of Economic Studies* 88.5, pp. 2275–2307.

Donovan, K., W. J. Lu, and T. Schoellman (2021). *Labor market dynamics and development*. Working Paper 596. Federal Reserve Bank of Minneapolis.

Emerick, K. (2018). "Agricultural productivity and the sectoral reallocation of labor in rural India". *Journal of Development Economics* 135, pp. 488–503.

Fields, G. S. (2009). "Segmented labor market models in developing countries". *The Oxford Handbook of Philosophy of Economics*. Ed. by D. Ross and H. Kincaid. Oxford University Press, pp. 476–507.

Glinskaya, E. E., A. I. De Kleine Feige, L. H. Vu Thi, G. T. Long, T. Hoang, C. N. Viet, B. R. Larson, S. Sen, R. J. Palacios, S. Bales, E. Masaki, and N. N. Nguyen (2021). *Vietnam - Adapting to an Aging Society (English)*. Report. World Bank Group.

Gollin, D. (2014). "The Lewis model: A 60-year retrospective". *Journal of Economic Perspectives* 28.3, pp. 71–88.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). "Agricultural productivity differences across countries". *American Economic Review* 104.5, pp. 165–70.

Graff Zivin, J. and M. Neidell (2014). "Temperature and the allocation of time: Implications for climate change". *Journal of Labor Economics* 32.1, pp. 1–26.

GSO and ILO (2018). *2016 Report on informal employment in Vietnam*. Tech. rep. Government Statistics Office of Vietnam and International Labor Organization.

Hamory, J., M. Kleemans, N. Y. Li, and E. Miguel (2021). "Reevaluating agricultural productivity gaps with longitudinal microdata". *Journal of the European Economic Association* 19.3, pp. 1522–1555.

Herrendorf, B., R. Rogerson, and Á. Valentinyi (2014). "Chapter 6: Growth and Structural Transformation". *Handbook of Economic Growth*. Ed. by P. Aghion and S. N. Durlauf. Vol. 2. Elsevier B.V., pp. 855–941.

Herrendorf, B. and T. Schoellman (2018). "Wages, human capital, and barriers to structural transformation". *American Economic Journal: Macroeconomics* 10.2, pp. 1–23.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al. (2020). "The ERA5 global reanalysis". *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049.

Hobijn, B., T. Schoellman, and A. Vindas (2019). *Structural transformation by cohort*. Working Paper. Arizona State University.

Hsiang, S. M. and R. E. Kopp (2018). "An economist's guide to climate change Science". *Journal of Economic Perspectives* 32.4, pp. 3–32.

IPCC (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.

Jann, B. (2021). "Relative distribution analysis in Stata". *Stata Journal* 21.4, pp. 885–951.

Jessoe, K., D. T. Manning, and J. E. Taylor (2018). "Climate change and labour allocation in rural Mexico: Evidence from annual fluctuations in weather". *Economic Journal* 128.608, pp. 230–261.

Kim, D.-I. and R. H. Topel (1995). "Labor Markets and Economic Growth: Lessons from Korea's Industrialization, 1970-1990". *Differences and Changes in Wage Structures*. Ed. by R. B. Freeman and L. F. Katz. University of Chicago Press, pp. 227–264.

Kongsamut, P., S. Rebelo, and D. Xie (2001). "Beyond balanced growth". *Review of Economic Studies* 68.4, pp. 869–882.

Kuznets, S. (1973). "Modern economic growth: findings and reflections". *American Economic Review* 63.3, pp. 247–258.

Lewis, W. A. (1954). "Economic development with unlimited supplies of labour". *Manchester School of Economic and Social Studies* 22, pp. 139–191.

Liu, M., Y. Shamdasani, and V. Taraz (forthcoming). "Climate change and labor reallocation: Evidence from six decades of the Indian Census". *American Economic Journal: Economic Policy*.

Liu, Y., C. B. Barrett, T. Pham, and W. Violette (2020). "The intertemporal evolution of agriculture and labor over a rapid structural transformation: Lessons from Vietnam". *Food Policy* 94, p. 101913.

Lowder, S. K., J. Skoet, and T. Raney (2016). "The number, size, and distribution of farms, smallholder farms, and family farms worldwide". *World Development* 87, pp. 16–29.

Maccini, S. and D. Yang (2009). "Under the weather: Health, schooling, and economic consequences of early-life rainfall". *American Economic Review* 99.3, pp. 1006–26.

McCaig, B. and N. Pavcnik (2015). "Informal employment in a growing and globalizing low-income country". *American Economic Review* 105.5, pp. 545–50.

— (2017). "Moving out of agriculture: Structural change in Vietnam". *Structural change, fundamentals, and growth: A framework and case studies*. Ed. by M. McMillan, D. Rodrik, and C. P. Sepulveda. Washington, D.C: International Food Policy Research Institute.

McCaig, B. and N. Pavcnik (2018). "Export Markets and Labor Allocation in a Low-Income Country". *American Economic Review* 108.7, pp. 1899–1941.

Nath, I. (2020). *The food problem and the aggregate productivity consequences of climate change*. NBER Working Paper w27297. National Bureau of Economic Research.

Ngai, L. R. and C. A. Pissarides (2007). "Structural change in a multisector model of growth". *American Economic Review* 97.1, pp. 429–443.

Ngai, L. R. and B. Petrongolo (2017). "Gender gaps and the rise of the service economy". *American Economic Journal: Macroeconomics* 9.4, pp. 1–44.

Nguyen, T. T. A., M. D. Luu, and D. C. Trinh (2016). "The evolution of Vietnamese industry". *Manufacturing transformation: comparative studies of industrial development in Africa and emerging Asia*. Ed. by J. Page, F. Tarp, J. Rand, A. Shimeles, C. Newman, and M. Söderbom. Oxford University Press.

Ortiz-Bobea, A. (2021). "The empirical analysis of climate change impacts and adaptation in agriculture". *Handbook of Agricultural Economics*. Vol. 5. Elsevier, pp. 3981–4073.

Park, R. J., J. Goodman, M. Hurwitz, and J. Smith (2020). "Heat and learning". *American Economic Journal: Economic Policy* 12.2, pp. 306–39.

Pham, T. (2022). "The child education and health ethnic inequality consequences of climate shocks in Vietnam". *Economics of Education Review* 90, p. 102311.

Porzio, T., F. Rossi, and G. Santangelo (2022). "The Human Side of Structural Transformation". *American Economic Review* 112.8, pp. 740–98.

Raymond, C. (2008). ""No Responsibility and No Rice": The Rise and Fall of Agricultural Collectivization in Vietnam". *Agricultural History* 18.1, pp. 43–61.

Raymond, C., T. Matthews, and R. M. Horton (2020). "The emergence of heat and humidity too severe for human tolerance". *Science Advances* 6.19, eaaw1838.

Rode, A., R. E. Baker, T. Carleton, A. D'Agostino, M. Delgado, T. Foreman, D. R. Gergel, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, A. Jina, R. E. Kopp, S. B. Malevich, K. E. McCusker, I. Nath, M. Pecenco, J. Rising, and J. Yuan (2022). *Labor disutility in a warmer world: The impact of climate change on the global workforce.* Working Paper.

Roy, A. D. (1951). "Some Thoughts on the Distribution of Earnings". *Oxford Economic Papers* 3.2, pp. 135–146.

Santangelo, G. (2019). *Firms and Farms: The Local Effects of Farm Income on Firms' Demand*. Cambridge Working Papers in Economics 1924. University of Cambridge.

Schlenker, W. and M. J. Roberts (2009). "Nonlinear temperature effects indicate severe damages to US crop yields under climate change". *Proceedings of the National Academy of Sciences* 106.37, pp. 15594–15598.

Schoellman, T. (2020). "Comment on "Migration costs and observational returns to migration in the developing world"". *Journal of Monetary Economics* 113, pp. 155–157.

Sherwood, S. C. and M. Huber (2010). "An adaptability limit to climate change due to heat stress". *Proceedings of the National Academy of Sciences* 107.21, pp. 9552–9555.

Somanathan, E., R. Somanathan, A. Sudarshan, and M. Tewari (2021). "The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing". *Journal of Political Economy* 129.6, pp. 1797–1827.

Turner, M. G., W. J. Calder, G. S. Cumming, T. P. Hughes, A. Jentsch, S. L. LaDeau, T. M. Lenton, B. N. Shuman, M. R. Turetsky, Z. Ratajczak, J. W. Williams, A. P. Williams, and S. R. Carpenter (2020). "Climate change, ecosystems and abrupt change: science priorities". *Philosophical Transactions of the Royal Society B* 375.1794, p. 20190105.

Ulyssea, G. (2018). "Firms, informality, and development: Theory and evidence from Brazil". *American Economic Review* 108.8, pp. 2015–47.

— (2020). "Informality: Causes and consequences for development". *Annual Review of Economics*.

Vasseur, D. A., J. P. DeLong, B. Gilbert, H. S. Greig, C. D. Harley, K. S. McCann, V. Savage, T. D. Tunney, and M. I. O'Connor (2014). "Increased temperature variation poses a greater risk to species than climate warming". *Proceedings of the Royal Society B: Biological Sciences* 281.1779, p. 20132612.

World Bank (n.d.). *Climate change knowlege portal: Vietnam*.

Zhang, Y., I. Held, and S. Fueglistaler (2021). "Projections of tropical heat stress constrained by atmospheric dynamics". *Nature Geoscience* 14.3, pp. 133–137.

Supplementary Materials
For Online Publication

Climate Change and Intersectoral Labor Reallocation in the Presence of Labor Market Frictions

Trinh Pham
Dyson School of Applied Economics and Management
Cornell University
tp347@cornell.edu

A Data and Measurement

A1 Data

A1.1 Employment and Labor Productivity Data

The household- and individual-level data are retrieved from the random 5% population and housing census in 1989 (Minnesota Population Center 2015), and the household livings standard survey conducted by the General Statistics Office of Vietnam (GSO) in 1993, 1998, and every two years since 2002 (GSO n.d.[a]). The household survey is representative at the national and provincial level.

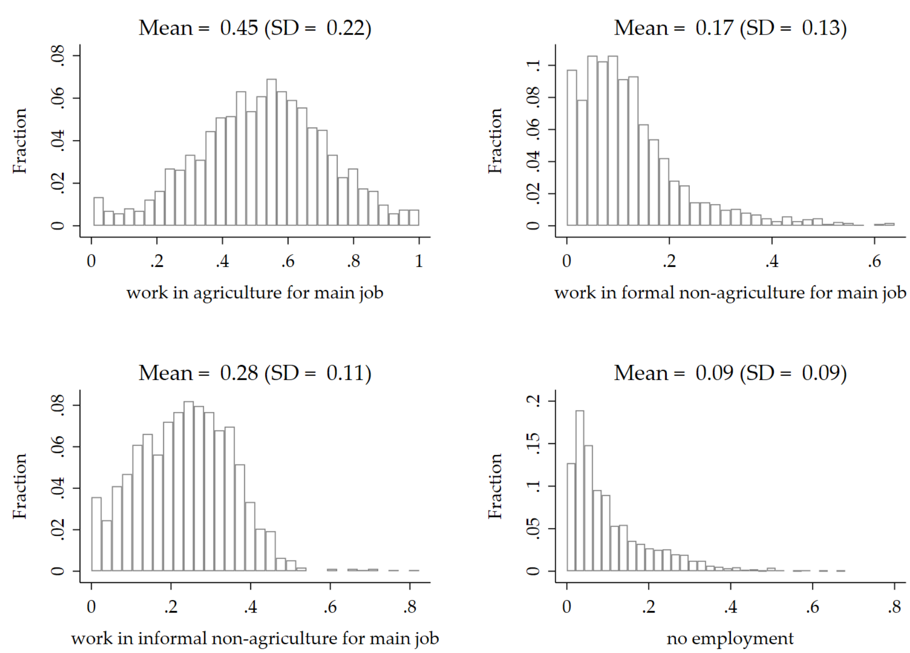
Table A1: List of Vietnamese Data Sets

Data Set	Reference Period	Sample size	Source	Data Access
Population and Housing Census 1989	12 months	5% census	IPUMS	Public
Living Standard Survey 1992/1993	12 months	4,800 households	GSO	Restricted
Living Standard Survey 1997/1998	12 months	6,000 households	GSO	Restricted
Household Living Standard Survey 2002	12 months	30,000 households	GSO	Restricted
Household Living Standard Survey 2004	12 months	45,000 households	GSO	Restricted
Household Living Standard Survey 2006	12 months	45,000 households	GSO	Restricted
Household Living Standard Survey 2008	12 months	45,000 households	GSO	Restricted
Household Living Standard Survey 2010	12 months	45,000 households	GSO	Restricted
Household Living Standard Survey 2012	12 months	45,000 households	GSO	Restricted
Household Living Standard Survey 2014	12 months	70,000 households	GSO	Restricted
Household Living Standard Survey 2016	12 months	45,000 households	GSO	Restricted
Household Living Standard Survey 2018	12 months	70,000 households	GSO	Restricted
Annual Enterprise Census 2002 to 2016	Fiscal year	all formal firms	GSO	Restricted

The key variable of interest is employment in agriculture, informal non-agriculture, and formal non-agriculture. The variable is constructed using data from the employment module of the survey, which covers hours worked, industries, as well as types of employer of the two most time-consuming jobs. I restrict the sample to 24-64 year old workers with information on industry of employment and types of employer to capture working-age individuals with completed education.

Province-level longitudinal employment dataset For the main temperature-sectoral employment analysis, I compute properly weighted share of individuals working their principal job in agriculture, informal and formal non-agriculture, for each of the three age groups (24-39 years old, 40-54 years old, and 55-64 years old) in each of the 52 provinces for each of the 11 survey waves over the study period. The final sample includes 1,707 province-age-year observations.[1] For the temperature-sectoral hours worked analysis, in 2002, only information of the principal job is collected. To ensure the measure of hours worked in a sector is consistent over time, I drop data from the survey wave 2002, which ends up having 1,551 province-age-year observations.

Figure A1: Summary Statistics on Province-level Employment Shares



Individual-level longitudinal dataset Although the household survey is repeated cross-sectional, it contains a (random) rotating panel sub-component that tracks households and individuals over a period of up to four years, which allows me to analyze individual transition from agriculture to informal and formal non-agricultural sectors over a longer time than is usually feasible. I link individuals over time using a unique individual identification code based on household identification, and

[1] In 1993 and 1998, only 50 and 51 provinces, respectively, were surveyed.

other individual information including gender, birth year, and sometimes confidential information (e.g., full name) provided by GSO in order to ensure the matching is correct.[2]

Province-level longitudinal production dataset I assemble a province-level production dataset separately for agriculture, informal non-agriculture, and formal non-agriculture.

Agriculture and Informal Non-Agriculture: I combine multiple waves of the household survey from 2002 to 2016 to construct household-level agricultural and informal non-agricultural production datasets.

The key variable of interest is annual revenue per worker. Agricultural revenues comprise of revenues from crops, livestock, aquaculture, forestry, and farm services. Informal non-agricultural revenues include revenues from non-farm business. The number of workers are measured as the number of household members engaging in agriculture and non-agriculture as their primary job. These information are reported by the households for the 12-month reference period before the interview. Correspondingly, I restrict the sample to households that do not hire labors, because the household survey does not record information on the number of hired workers.

Household-level data are then merged with weather data in the same province using the timing of interview, similar to employment data. I then aggregate household-level to provincial-level dataset by taking a weighted average of all household producers in that province, with weight being the production size (i.e., number of workers).

Formal Non-Agriculture: The firm-level data are retrieved from the annual census conducted by the General Statistics Office of Vietnam since 2001 (VEC) (GSO n.d.[b]). While the household survey has the advantage of covering both formal and informal workers but the shortcoming that it is at best representative at the province level, VEC has the advantage of a census and being available at yearly level, but small and informal firms are not covered.

The enterprise census collects rich information on ownership type, industry

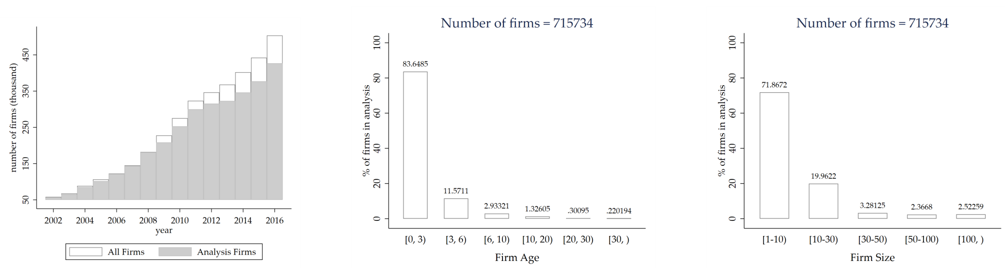
[2] The matching codes for the survey waves 2002 to 2006, and 2010 to 2012 are graciously shared by McCaig and Pavcnik (2015) and McCaig and Pavcnik (2021), respectively.

type, employment, labor compensations, as well as business performance and financial information of registered firms in the preceding fiscal year. I construct a dataset from 2002-2016 for firms whose main economic activity is non-agriculture using a unique firm identification code which comprises of tax code (available all periods), firm code (available before 2012), and branch code (available before 2014). New firms that do not have tax code yet are identified by a unique firm code assigned by the survey team.

The key variable of interest is annual revenue per worker, where revenue is calculated as the net turnover of goods and services, and the number of workers are measured at the end of fiscal year. In constructing the production data, I impose the following conditions: (i) The firm should not operate more than one branch, (ii) The firm should be in operation and report positive revenues, (iii) The firm should report positive number of workers at year end. Restriction (i) drops 0.7% of the original sample. Restrictions (ii) and (iii) mainly reflect data errors, and drop 9.6% and 0.008%, respectively, of the original sample. Firm data are then merged with weather data in the same province. I then aggregate firm-level to provincial-level dataset by taking a weighted average of all firms in that province, with weight being the firm size.

Panel A of Figure A2 shows the number of firms in the analysis sample, which reflects the increasing number of registered firms in Vietnam over the same period. As in many other low and middle income countries, a majority of these firms are young and small: more than 80% are less than three years old (Panel B), and nearly 70% have fewer than 10 workers (Panel C). Firms operating more than 10 years in the market account for less than 3% of the total sample.

Figure A2: Non-agricultural firm-level data



(a) Number of Firms (b) Age Distribution (c) Size Distribution

A1.2 Weather Data

This section summarizes how weather variables are constructed from the ERA5 reanalysis data.

Wet-bulb temperature Wet-bulb temperature (WBT) is a nonlinear function of dry-bulb temperature (i.e., ambient air temperature) and relative humidity. It reflects the lowest temperature to which air can be cooled by the evaporation of water into the air at a constant pressure. The measure of WBT has been increasingly used in the economics literature to study the combined effects of heat and humidity on worker productivity (Adhvaryu, Kala, and Nyshadham 2020; Somanathan et al. 2021; LoPalo forthcoming).

To calculate daily average WBT, I proceed in three steps. First, I calculate hourly relative humidity RH, which is defined as the ratio of vapor pressure e and saturation vapor pressure e_s, using hourly air temperature T_a (°C) and hourly dew-point temperature T_d (°C), following Bolton (1980):[3]

$$RH = 100 \times \frac{e}{e_s} = 100 \times \exp\left[\frac{17.67 \times 243.5 \times (\mathrm{T}_d - \mathrm{T}_a)}{(243.5 + \mathrm{T}_a)(243.5 + \mathrm{T}_d)}\right]$$

Second, I calculate hourly WBT using hourly dry-bulb temperature T_a (°C) and hourly relative humidity RH (%), following Stull (2011):

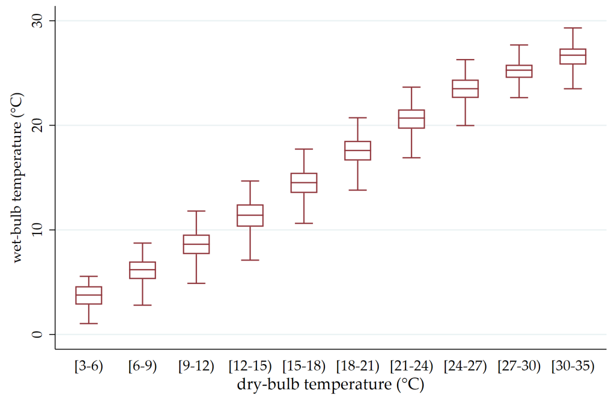
$$\begin{aligned} \mathrm{WBT} = \quad & \mathrm{T}_a \times atan\left[0.151977 \times (RH + 8.313659)^{0.5}\right] + atan(\mathrm{T}_a + RH) - 4.686035 \\ & - atan(RH - 1.676331) + 0.00391838(RH)^{1.5} \times atan(0.023101 \times RH) \end{aligned}$$

Finally, I take the mean of hourly T_w to get daily WBT. Figure A3 plots the relationship of the two measures of temperature. At any given dry-bulb temperature level, there is a large variation in WBT. For example, for the 27-30°C dry-bulb temperature interval, WBT ranges from 18 to 28°C.

Precipitation Daily precipitation is calculated as the sum of hourly precipitation. I then compute the second order polynomial of daily precipitation at each grid-level. This is done before the data are spatially averaged in order to accurately represent the distributions at grid level.

[3] This equation is available as archive on University Corporation for Atmospheric Research's website.

Figure A3: Dry-bulb temperature and wet-bulb temperature (°C), 1980-2020



Notes: This figure displays the median, middle 50%, and the adjacent values of daily mean dry bulb temperatures associated with various daily mean wet bulb temperatures.

Extreme precipitation I also construct standardized precipitation index for each month/year as the deviation of the observed precipitation from the long-term mean divided by the historical standard deviation:

$$SPI_{pmy} = \frac{R_{pmy} - \bar{R}_{pm}}{\sigma_{pm}}$$

where R_{pmy} is the observed rainfall for a given month m of year y in province p. $\bar{R}_{pm}$ is the long-term mean rainfall in province p in month m over the 30-year period 1990-2020. σ_{pm} is the corresponding standard deviation. The index helps determine the level of excess relative to the climatological norm for the location. A province is considered having excess rainfall in month m of year y relative to the long-term mean if its $SPI_{pmy} \geq 1$.

Wind speed The data include wind components, which are eastward and northward wind vectors, represented by the variables "U" and "V" respectively. The U wind component is parallel to the x-axis (i.e., longitude) with a positive (negative) U wind coming from the west (east). The V wind component is parallel to the y-axis (i.e., latitude) with a positive (negative) V wind coming from the south (north).

I calculate hourly value of wind speed, which is the magnitude of the wind vector, using hourly U and V components according to the Pythagorean Theorem:[4]

[4]GES DISC Data in Action: Calculate Wind Speed and Direction using U and V components.

wind speed = $\sqrt{U \times U + V \times V}$. Daily wind speed is then calculated by taking the maximum value of hourly wind speed for the corresponding day.

Aggregation of grid-level weather data to province-level weather data I transform grid-level weather data to province-level weather data using two methods. The first method is to take weighted average of four nearest grid points to province centroids, where the weight is inverse distance. The second method is to average all the points within the geographic boundary of the first administrative level–a province, except for wind speed where I use a maximum value. In both cases, non-linear transformations of temperature and rainfall are computed at the grid level before averaging values across space, and finally summing over days during the reference period. This procedure is similar to Carleton et al. (2022).

To see how this calculation is conducted, consider the fourth-order polynomial specification for temperature. I begin with data on average temperatures for each day d at each grid point z, generating observations WBT_{zd} These grid-level values are aggregated to the province level p for each 12-month reference period. To do this, I first raise grid-level temperature to the power n, computing $(\text{WBT}_{zd})^n$ for $n \in \{1, 2, 3, 4\}$. I then take a spatial average of these values following the two methods mentioned above. I then sum these daily polynomial terms $(\text{WBT}_{zd})^n$ over days during individual-specific reference periods, i.e., 12 full months before the survey interview. This nonlinear transformation performed prior to aggregation allows the aggregated measure of temperature to capture grid-by-day level exposure to very hot and very cold temperatures. Quadratic polynomials in precipitation are similarly calculated.

Because there has been changes in the administrative boundaries in Vietnam over the last three decades and most of the changes happen in case of splitting, I use the original administrative units in 1993, which gives a consistent sample of 52 provinces over the study period. An exception is that Ha Tay province was merged into Hanoi city in 2008 and thus I use the boundary of the new Ha Noi for consistency. This process results in the province-level vector of weather data for each 12-month interval.

A1.3 Other Data

I assemble a longitudinal dataset of yields for two major crops including rice and maize at the province level from 1998 to 2018. The data are then collapsed into the consistent provincial level during the study period. The analysis panel consists of 52 provinces over 10 years, biennially from 1998 to 2018. I also construct a panel dataset of in-migration, out-migration, and net-migration rates at the province level covering every two years from 2008 to 2018.[5]

A2 Informality Measurement

Informality can broadly be defined either from the worker side or from the employer side. According to GSO and ILO (2018), informality on the worker side implies that workers do not have social security benefits and labor contract with a minimum term of three months ("informal workers"). On the employer side, informality implies that firms do not have legal status or register with the government ("informal firms"). In this paper, the notion of informality largely follows that from the employer side by assuming individuals who either are self-employed or work for household businesses or collectives are employed informally ("informal employment").

The household surveys include a question on whether a worker has benefited from social insurance since the 2010 wave, and a question on whether she has a labor contract since the 2014 wave. Although these two questions do not perfectly capture the definition of informal workers defined by GSO and ILO (2018), they allow a cross-check between the definition of informality employed in this paper and of informal workers in 2014, 2016, and 2018.

Table A2 shows that the two definitions are largely similar. The Pearson correlation coefficient between informal workers and informal employment variables is nearly 0.9. Only a small fraction (less than 0.2%) of formal workers are classified as informally employed. The notion of informal employment, however, does not capture very well the intensive margin of informality: up to 14.5% of workers in formal firms does not have social security benefits and labor contract. Table A3 provides further details by industry and highlights the differences in education, as proxied by years of schooling, between workers in the informal and formal sectors across industries. Generally, workers in the informal sector have lower educational attainment compared to their peers in the formal non-agricultural sector.

[5] These data are available on the GSO website, at https://www.gso.gov.vn/en/statistical-data. Because the agricultural data are available from 1995, while the migration data are only available from the years 2005/2007 onward, I restrict the analysis data to those years overlapping with the study period.

Table A2: Informality

	(1) Pearson Correlation Coefficient between Informal Workers and Informal Employment	(2) Share of Formal Workers in Informal Employment	(3) Share of Informal Workers in Formal Employment
2014	0.8842	0.0021	0.1451
2016	0.8864	0.0017	0.1429
2018	0.8917	0.0021	0.1288
Total	0.8875	0.0020	0.1389

NOTES: Informal employment are defined as self-employment, employment in household businesses, and collectives. Informal workers are defined as those who do not have social security benefits nor labor contract. *Source:* Data from VHLSS 2014, 2016, 2018.

Table A3: Education level of workers by industry and informality, non-agricultural sectors

	Years of Schooling		Share of Informal Workers in Formal Firms		
	(1)	(2)	(3)	(4)	(5)
	Informal	Formal	2014	2016	2018
Manufacturing: high tech	8.945	11.086	0.037	0.015	0.031
Manufacturing: medium tech	8.192	10.686	0.159	0.117	0.125
Manufacturing: low tech	7.868	9.493	0.135	0.119	0.103
Service: knowledge intensive	9.440	13.296	0.060	0.061	0.050
Service: less knowledge-intensive	7.982	11.290	0.264	0.281	0.226
Mining and quarrying	7.405	10.831	0.131	0.140	0.147
Public utilities	8.737	11.884	0.053	0.077	0.073
Construction	7.805	10.323	0.474	0.466	0.473
Total	8.297	11.111	0.164	0.160	0.154

A3 Tradability Measurement

To construct the tradability measurement based on agricultural price, I proceed in four steps.

(i) The household survey has a module on agricultural production, which collects information on the monetary value and amount at harvest of all crops grown by a household. I compute price as total monetary value divided by total amount of harvest for each crop.

(ii) I assemble a dataset of household-level price at harvest for three types of rice: winter-spring ordinary rice, summer-autumn ordinary rice, and autumn-winter (Mua) ordinary rice. Based on province-specific agricultural rice production calendar, I then assign each price to the corresponding harvesting month, for example, price at harvest of winter-spring rice is the June rice price in Red River Delta region.[6]

(iii) I aggregate household-level price to province-level price by taking the median price of all households in the same province. The price in nominal Vietnamese Dong is then converted to nominal US Dollar using World Bank's official exchange rates.[7] The analysis dataset consists of 52 price time series for 52 provinces.

(iv) Such local rice monthly price time series are then compared with the monthly world market price of Vietnamese rice 5% broken to obtain pairwise correlation coefficients.[8]

[6]These information are computed based on region-specific planting months reported by the Vietnam Academy of Agricultural Sciences (VAAS). More details in Vietnamese at https://vaas.vn/kienthuc/Caylua/01/index.htm.

[7]The exchange rate data are available at https://data.worldbank.org/indicator/PA.NUS.FCRF?locations=VN.

[8]The data on monthly world market price of Vietnamese rice 5% broken are obtained from the World Bank's "Pink Sheet," available at https://www.worldbank.org/en/research/commodity-markets.

B Robustness Checks and Additional Results using Dry-bulb Temperature

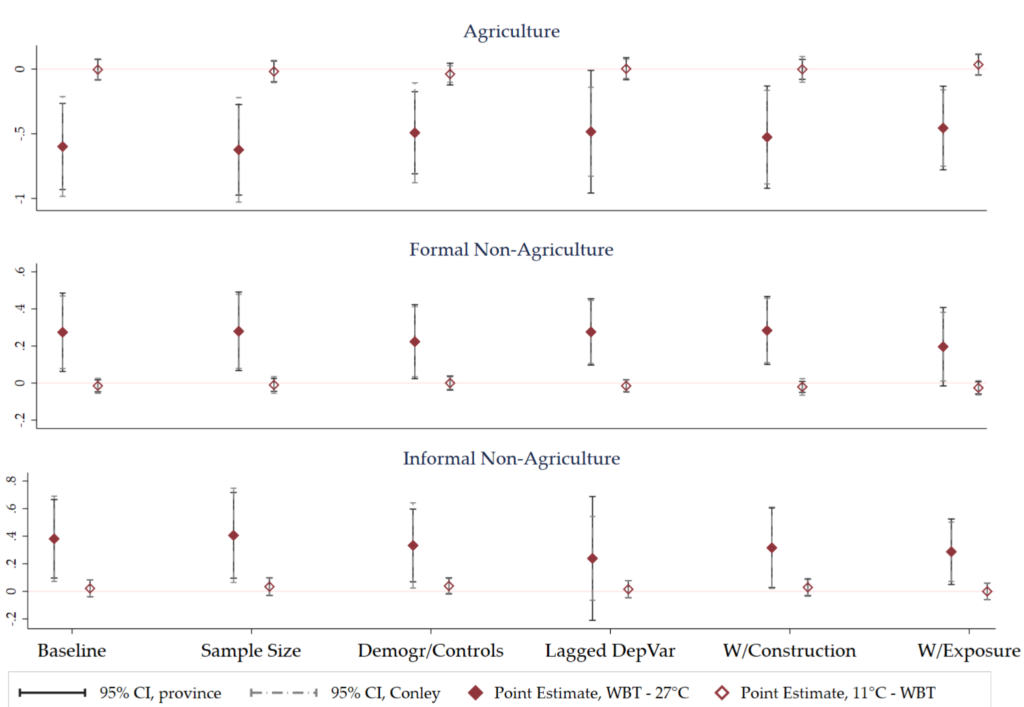
B1 Robustness Checks

Average Effects Figures B1 and B2 present the key coefficients on the average temperature effects using the panel approach (equation 1), and long differences approach (equation 3), respectively.

First, I test whether the results are robust to changes in the sample. In the main analysis, I keep all province-age group-year cells which were constructed using less than 30 individual observations. These cells are mostly from the 1993 and 1998 rounds of the household survey where the sample is relatively small with approximately 5,000 households nationwide. I report results using the subset of cells which were constructed with at least 30 individual observations. The effects are largely unchanged under the exclusion of those cells.

Figure B1: Robustness Check: Wet-bulb Temperature and Sectoral Employment

The Effects of Cumulative WBT: Panel Approach



A particular concern over examining the effects of temperature on sectoral labor allocation is that education effects might confound the temperature effects. The country's extensive education expansion over the last few decades (Dang and Glewwe 2018) might have equipped individuals with skills that are more valuable in non-agricultural sectors–the leading explanation for the essential role that new birth cohorts play over the course of structural transformation (Porzio, Rossi, and Santangelo 2022). I test this concern by controlling for time-varying demographic characteristics, including educational attainment, share of male workers, and share of Kinh ethnic majority when estimating equations (1) and (3). The inclusion of such variables could help absorb residual variation and produce more precise estimates but could also be problematic if these variables themselves are considered outcome variables. The coefficients on sectoral employment shares are of slightly smaller magnitude to those obtained from the baseline specification but remain statistically significant, suggesting that changes in these demographic characteristics cannot explain for the changes in sectoral employment shares.

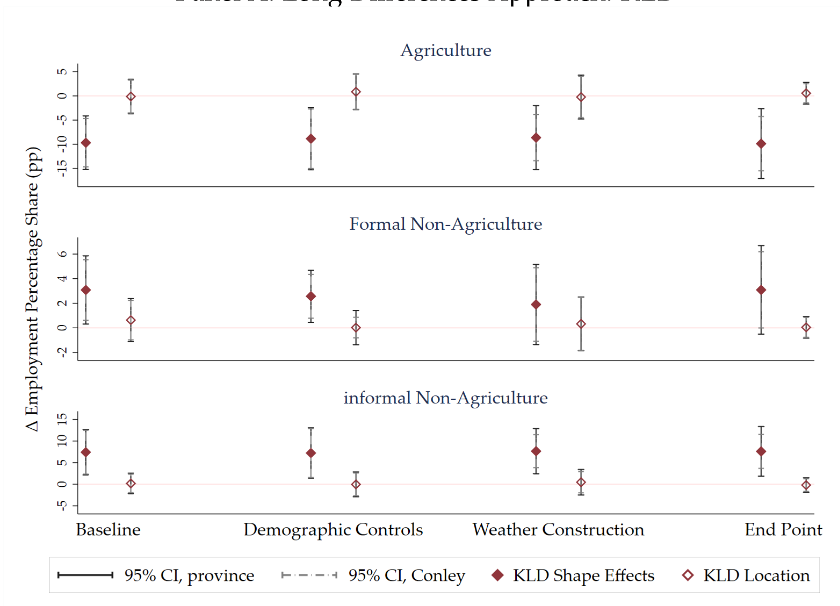
Third, because sectoral employment shares at the local level are highly correlated from one year to the next, estimation from equation (1) might suffer from omitted variable bias. I explore this concern by controlling for the lagged value of the dependent variable in the preceding period. A drawback of estimating this dynamic panel model is that it is inconsistent when lagged dependent variables and fixed effects are estimated simultaneously with OLS (Nickell 1981). This concern is especially prominent when the length of the data panel is short. As seen in Figure B1, the coefficients obtained from such a dynamic panel model are generally of similar magnitude to the coefficients from the baseline specification. The effect on informal non-agricultural labor share is less precisely estimated and of smaller magnitude relative to the baseline results, however.

Fourth, I test the robustness of the results to an alternative construction of the weather measures. In the baseline analysis, provincial level weather variables are computed as the weighted average of the four grid points closest to provincial geographic centroids, with weights being the inverse distance of weather grids to province centroids. The results obtained from both panel and long differences approaches also hold under an alternate construction where province-level weather variables are computed as the average value of all grid points within the geographical boundary of a province.

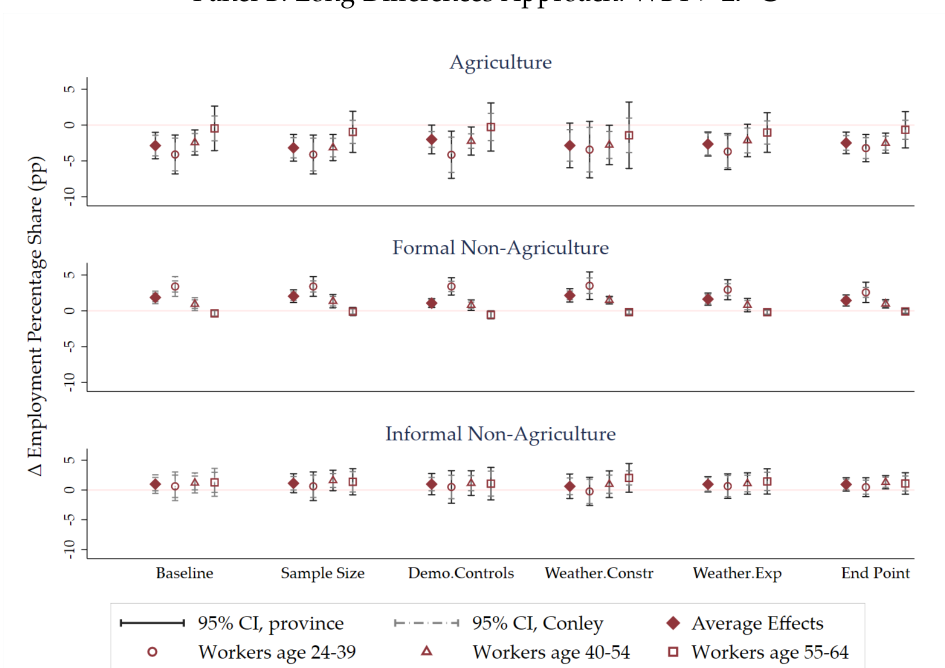
Fifth, in merging the weather data with individual-level data, I assume that

Figure B2: Robustness Check: Wet-bulb Temperature and Sectoral Employment

Panel A: Long Differences Approach: KLD



Panel B: Long Differences Approach: WBT > 27°C



NOTES: This figure presents the effects of temperature change, proxied by KLD shape and location differences (Panel A) and hot temperatures (Panel B), on sectoral employment shares, which are obtained from estimating equation (3). Province distances are computed from province geographic centroids.

individuals were exposed to the weather distribution of the full 12 months prior to the timing of survey interview. I address the possibility that temperatures can exhibit lagged effects on decision to switch sector by using an alternate exposure: I assign to each individual the weather distribution of the full 14 months before the survey time.[9] The new estimates are of smaller magnitude to the corresponding baseline coefficients but remain statistically significant.

Sixth, I test the robustness of the long differences results to an alternative end point. In the baseline specification, I take the difference in outcomes (and weather variables) between two periods: 1992-2006 and 2007-2018. The first period covering 14 years comprises of five waves of the household survey 1993, 1998, 2002, 2004, and 2006. The second period covering 12 years comprises of six survey waves 2008, 2010, 2012, 2014, 2016, and 2018. In an alternate estimation, I divide the study period into two sub-periods: 1992-2004 and 2005-2018. The effects on sectoral employment shares are of similar magnitudes and remain statistical significance at conventional levels in the specification with KLD measures. In the specification with extreme temperature measures, while the coefficients are of similar signs to the baseline specifications, they are less precisely estimated, with formal non-agriculture results being an exception.

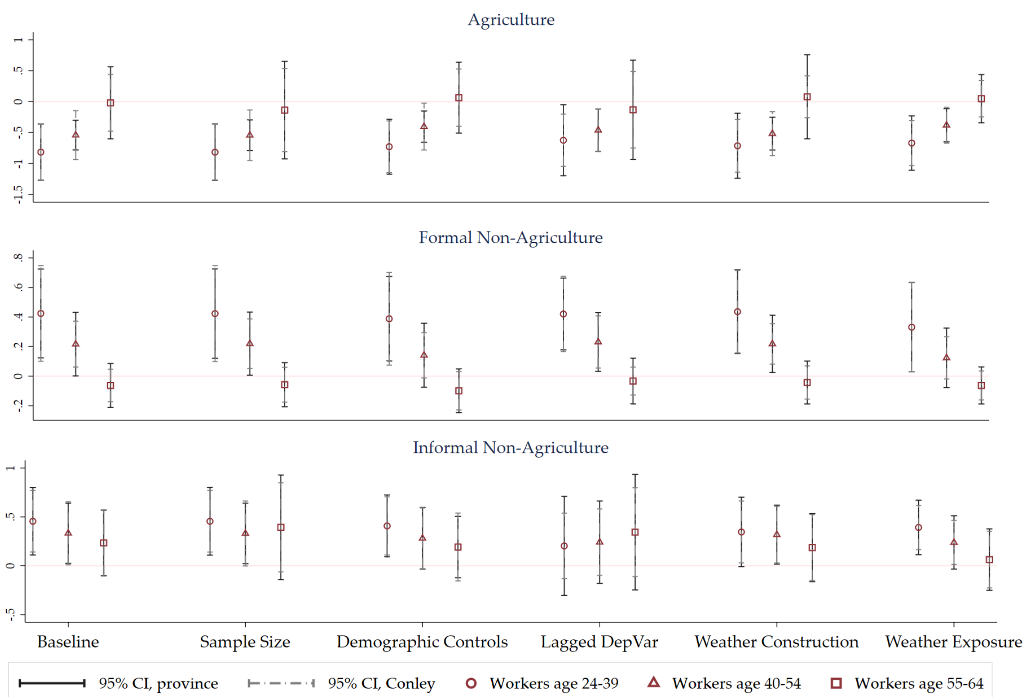
Heterogeneous Effects by Age Group Additional robustness check results for age group heterogeneity analysis using the panel approach are available in Figure B3. Consistent with findings from the baseline specification, the effect of hot days on agricultural labor shares declines as one moves from the youngest to the oldest group. The corresponding effect on employment shares in formal non-agriculture also decreases as one moves from the youngest to the oldest groups. As for informal non-agricultural labor shares, however, the effect of hot temperatures is similar across the three groups.

Figure B4 likewise plots the predicted sectoral-employment share-temperature response function that varies across age groups, estimated from models of fourth-order polynomials of daily average wet-bulb temperature, and cumulative wet-bulb temperature bins. As seen, the heterogeneous effects across age groups are also robust to alternate functional forms of temperature.

[9] The results are also robust to other exposures such as 18 months.

Figure B3: Robustness Check: Wet-bulb Temperature and Sectoral Employment by Age Group

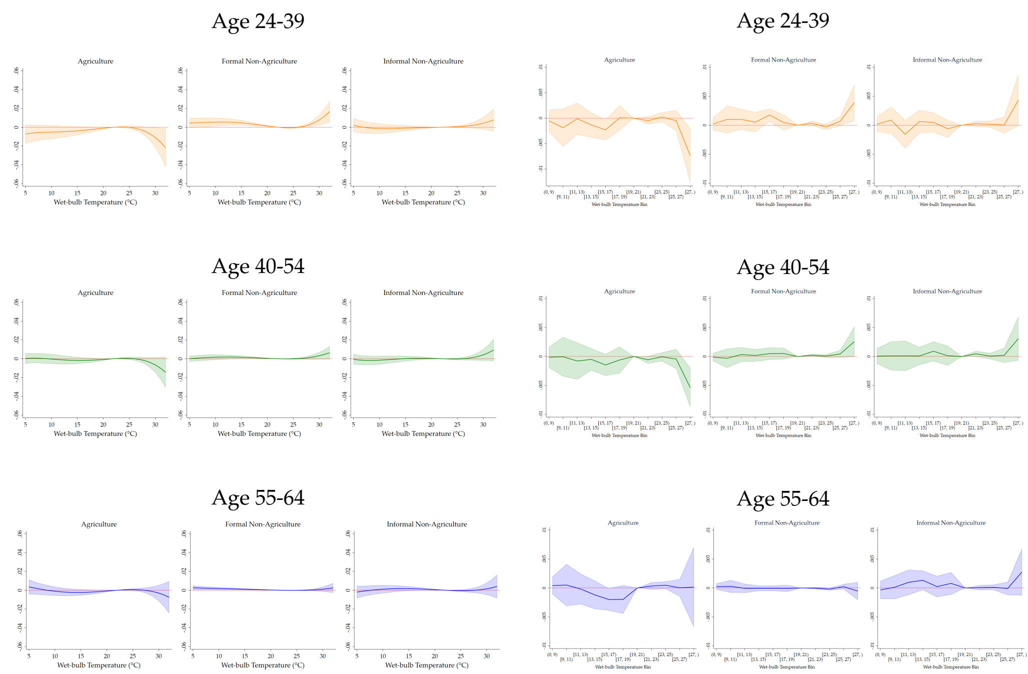
The Effects of Cumulative WBT above 27°C: Panel Approach



NOTES: Results from estimating equation (1). Dependent variables are percentage shares of employment in each sector. Province distances are computed from province geographic centroids. All regressions use sampling weights.

Figure B4: Robustness: Wet-bulb Temperature and Primary Sectoral Employment

Panel A: Fourth-order polynomials Panel B: Cumulative temperature bins

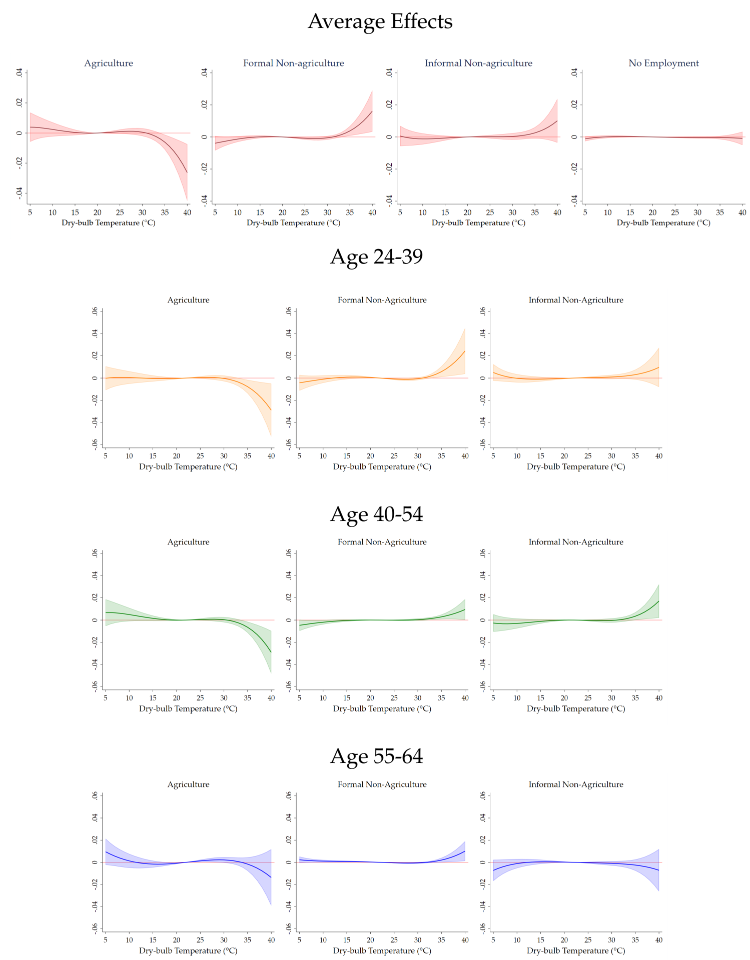


NOTES: Each graph represents a predicted sectoral employment share-temperature response function, estimated with equation (1). Shaded areas are 95% confidence interval. In Panel A, regression estimates are from a fourth-order polynomial in daily average temperature. In Panel B, regression estimates are from a model of cumulative temperature bins. All age-specific response functions for a sector are estimated jointly in a stacked regression model that is fully saturated with age group-specific fixed effects. Robust standard errors are clustered at the province level.

B2 The Effects of Dry-bulb Temperature on Sectoral Employment

Consistent with the findings using wet-bulb temperatures, high air temperatures cause a decrease in agricultural employment share, increases in formal and informal non-agricultural employment shares, but do not affect non-employment (Figure B5).

Figure B5: Dry-bulb Temperature and Sectoral Employment



NOTES: This figure represents a predicted sectoral employment share-temperature response function, estimated with equation (1), where estimates are from a fourth-order polynomial in daily average dry-bulb temperature. Shaded areas are 95% confidence interval. All age-specific response functions for a sector are estimated jointly in a stacked regression model that is fully saturated with age group-specific fixed effects. Robust standard errors are clustered at the province level.

C Additional Tables and Figures

Table C1: Wet-bulb Temperature and Migration Responses

	In Migration		Out Migration		Net Migration	
	(1)	(2)	(3)	(4)	(5)	(6)
(11 – WBT)	0.012	0.015	0.014	-0.028	-0.003	0.043
	(0.007)	(0.016)	(0.022)	(0.052)	(0.025)	(0.052)
.	[0.003]	[0.014]	[0.022]	[0.041]	[0.022]	[0.041]
(WBT – 27)	0.060	0.062	-0.081	-0.036	0.143	0.099
	(0.054)	(0.056)	(0.075)	(0.094)	(0.091)	(0.115)
.	[0.042]	[0.047]	[0.073]	[0.081]	[0.080]	[0.099]
N	312	312	312	312	312	312
Sample Mean of DepVar	5.06	5.06	6.71	6.71	-1.65	-1.65
Province FE	x	x	x	x	x	x
Region × Year FE	x	x	x	x	x	x
Province Linear Trend		x		x		x

NOTES: Unit of analysis is province-year. Dependent variables are migration rates (%) at the province level. All columns control for the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table C2: Wet-bulb Temperature and Agricultural Planting Area

	Rice (All)		Rice (Winter-Spring)		Maize		Grains	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(11 – WBT)	-0.024	-0.005	-0.001	0.003	-0.085	-0.065	-0.104	-0.065
	(0.025)	(0.019)	(0.011)	(0.012)	(0.087)	(0.080)	(0.085)	(0.076)
.	[0.031]	[0.013]	[0.012]	[0.008]	[0.064]	[0.046]	[0.075]	[0.047]
(WBT – 27)	-0.469	-0.248	0.024	0.019	0.268	0.102	-0.170	-0.115
	(0.385)	(0.151)	(0.102)	(0.034)	(0.168)	(0.117)	(0.403)	(0.184)
.	[0.436]	[0.147]	[0.080]	[0.050]	[0.263]	[0.123]	[0.571]	[0.188]
N	520	520	520	520	520	520	520	520
Sample Mean of DepVar	145.07	145.07	58.39	58.39	19.76	19.76	164.91	164.91
Province FE	x	x	x	x	x	x	x	x
Region × Year FE	x	x	x	x	x	x	x	x
Province Linear Trend		x		x		x		x

NOTES: Unit of analysis is province-year. Dependent variables are planting area, measured in thousand hectares, at the province level. All columns control for the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table C3: Wet-bulb Temperature and Crop Yields

	Rice (All)		Rice (Winter-Spring)		Maize		Grains	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(11 – WBT)	-0.000	-0.000	-0.001	-0.001	0.001	0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
.	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
(WBT – 27)	-0.020	-0.018	-0.040	-0.034	-0.009	0.005	-0.019	-0.016
	(0.008)	(0.007)	(0.013)	(0.012)	(0.010)	(0.007)	(0.008)	(0.007)
.	[0.009]	[0.007]	[0.016]	[0.010]	[0.008]	[0.006]	[0.009]	[0.007]
N	520	520	520	520	515	515	520	520
Sample Mean of DepVar	4.81	4.81	5.53	5.53	3.90	3.90	4.67	4.67
Province FE	x	x	x	x	x	x	x	x
Region × Year FE	x	x	x	x	x	x	x	x
Province Linear Trend		x		x		x		x

NOTES: Unit of analysis is province-year. Dependent variables are crop yields, measured in tonnes per hectare, at the province level. All columns control for the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 200 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table C4: Wet-bulb Temperature and Firm-Level Labor Productivity

	Mining and Quarrying		Construction	
	(1)	(2)	(3)	(4)
(WBT - 27)	-0.004439	-0.007775	-0.001341	0.000275
	(0.004822)	(0.005621)	(0.002934)	(0.002838)
(WBT - 27) × (Firm Size < 30 workers)	-0.011969		-0.000764	
	(0.003634)		(0.001877)	
(WBT - 27) × (Firm Age >= 10 years old)		-0.008676		-0.006184
		(0.002828)		(0.000931)
Total Temperature Effects (Small/Old Firms)	-0.016408	-0.016451	-0.002104	-0.005909
	(0.005597)	(0.005589)	(0.002881)	(0.002945)
Sample Mean of DepVar	4.77	4.77	4.88	4.88
Firm FE	x	x	x	x
Region-by-Year FE	x	x	x	x
Observation	23896	23896	398851	398851

NOTES: Unit of analysis is firm-year. Dependent variables are log of annual revenue per worker, measured in 2010 million VND. All columns control for firm age and firm size category dummies, cold temperatures (WBT < 11°C), the second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure and their interactions with firm category dummies. Robust standard errors clustered at the province level are in parentheses.

Table C5: Hot Wet-bulb Temperatures and Primary Sectoral Employment, by Gender and Education

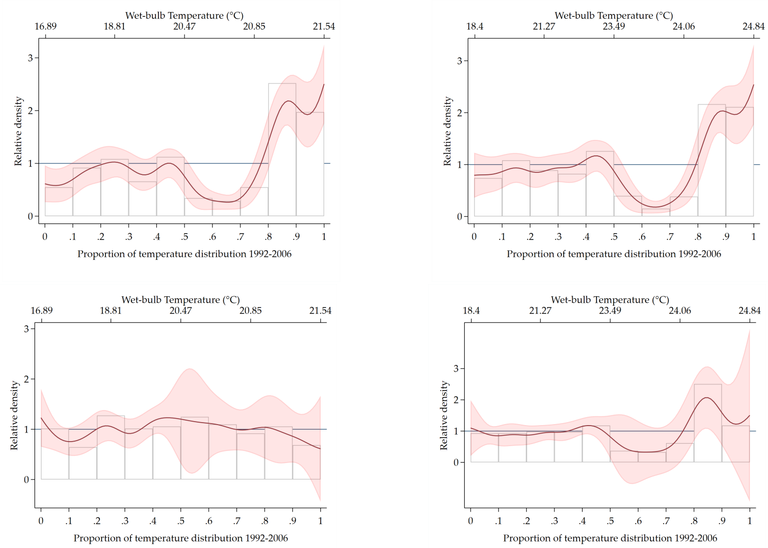
	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture	No Employment
	(1)	(2)	(3)	(4)
Male - Female	-0.021	-0.145	0.099	0.063
	(0.191)	(0.118)	(0.188)	(0.062)
Observations	1138	1138	1138	1138
Province × Gender FE	x	x	x	x
Region × Gender × Year FE	x	x	x	x
High School and Above - Below	0.325	0.069	-0.528	0.129
	(0.301)	(0.157)	(0.218)	(0.089)
Sample Mean of DepVar	45	17	28	9
Observations	1134	1134	1134	1134
Province × Education FE	x	x	x	x
Region × Education × Year FE	x	x	x	x

NOTES: This table presents the difference in the effects of hot temperatures (wet-bulb temperature degree days above 27°C) on sectoral employment shares between male and female workers, and between individuals who have a high school diploma or above and those who do not have a high school diploma. Unit of analysis is province-gender-year or province-education-year, respectively. Each cell is from a separate regression. All regressions use weights. Standard errors clustered at the province level in parentheses.

Figure C1: Change in wet-bulb temperature distribution in two provinces: 1992-2006 vs. 2007-2018

Location-unadjusted

Location-adjusted



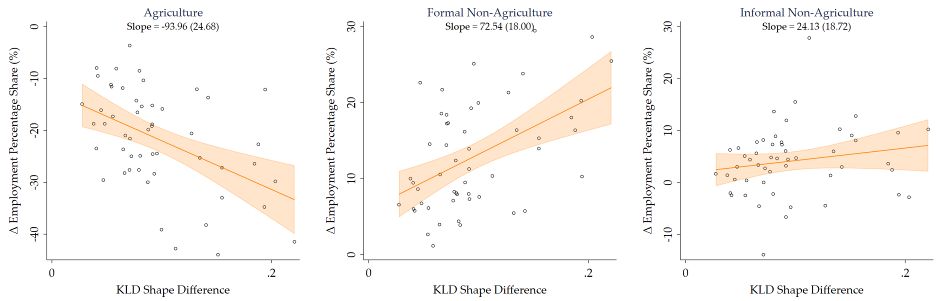
(a) Lam Dong
overall divergence = 0.26
location = 0.22, shape = 0.05
Δmean temperature = 0.31°C

(b) Binh Dinh
overall divergence = 0.26
location = 0.07, shape = 0.19
Δmean temperature = 0.33°C

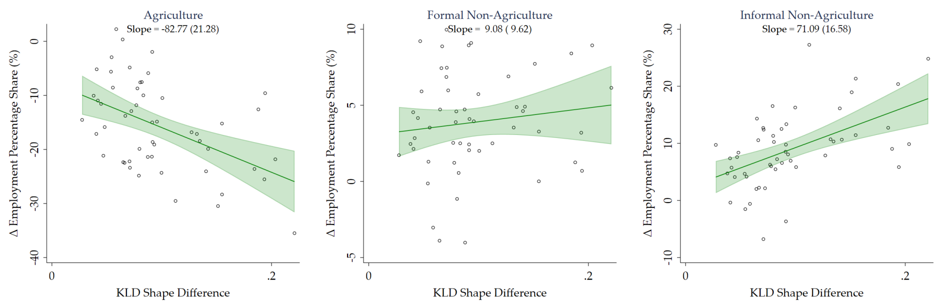
NOTES: This figure plots the relative density of the recent temperature distribution 2007-2018 relative to the reference temperature distribution 1992-2006, and 95% confidence interval from a non-parametric estimation using Epanechnikov kernel function with a bandwidth of 0.05 and 200 bootstraps. A relative density larger (smaller) than one means the recent distribution is overrepresented (underrepresented) relative to 1992-2006 at the corresponding level of temperature denoted on the top axis. While Lam Dong experiences a relatively smooth righward shift in the whole temperature distribution (location effect), Binh Dinh observes a polarization of temperature distribution with fewer mild days and more hot days in recent years (shape effect).

Figure C2: Change in temperature distribution and sectoral employment shares by age group

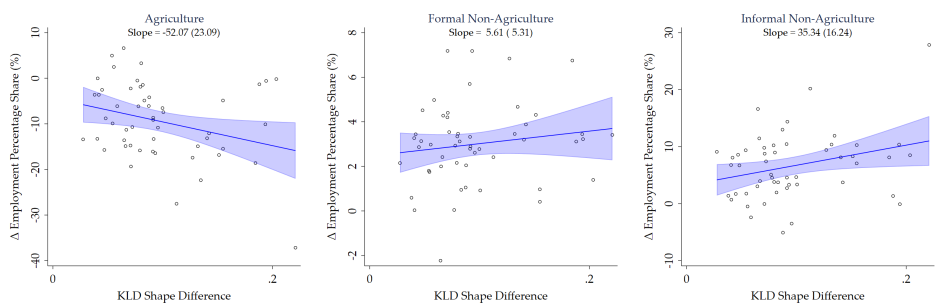
Panel A: Age 24-39



Panel B: Age 40-54



Panel C: Age 55-64



NOTES: Each panel shows the line of best fit and 95% confidence interval from the regression of change in sectoral employment percentage shares and change in temperature distribution, as proxied by shape difference measure, between the two periods: 1992-2006 and 2007-2018. Each circle represents an observation.

Figure C3: Distribution of Days by Average Wet-bulb Temperature Bin 1992-2018

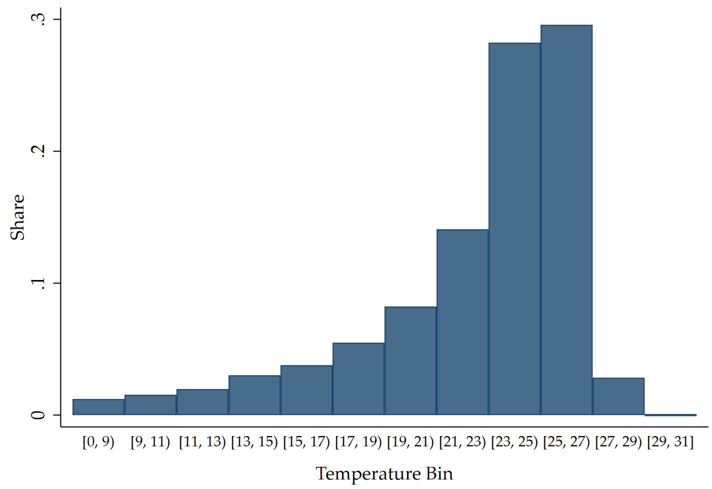


Figure C4: Interview Month of an Average Person in each Province: 1992-2018

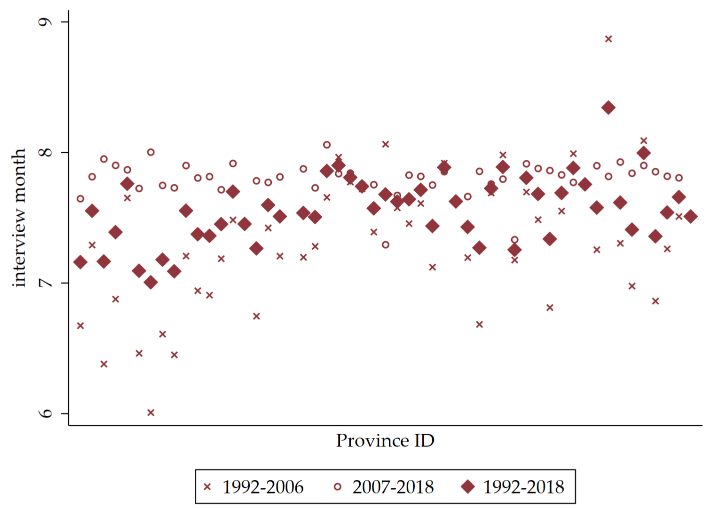
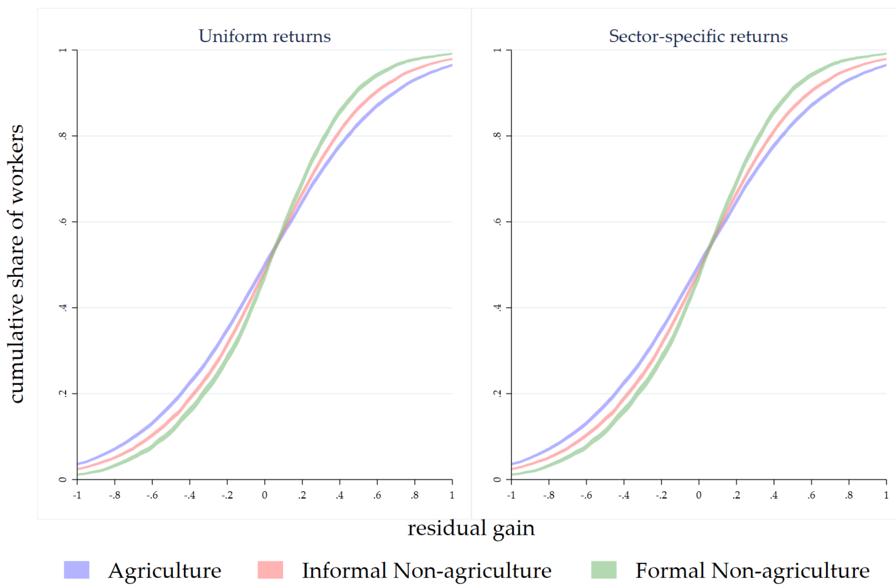


Figure C5: First-order stochastic dominance among sector-specific residualized distributions



Notes: Using the DASP packages developed by Araar and Duclos (2007), this figure plots the CDF (with 95% confidence interval) of the sector-specific residualized distributions obtained from estimating sector gaps in earnings using individual-level panel datasets for the sample of switchers. Under both assumptions of uniform and nonuniform returns to individual characteristics across sectors, there is no full first-order dominance between sector-specific residual gains. However, below the zero residual gain, both informal and formal non-agriculture first-order dominates agriculture, indicating that a smaller percentage of workers have negative gains if transitioning into non-agriculture. Similarly, above the zero residual gain, agriculture first-order dominates non-agriculture, indicating that a larger percentage of workers have positive gains if transitioning into non-agriculture. *Sources*: Data from five VHLSS three-wave individual panels 2002-2004-2006 to 2014-2016-2018.

D Estimating Temperature Effects on Marginal Product of Labor

Consider a firm or household's production function technology that can be represented by a production function $h(.)$ that relates output (Y), inputs $X = \left[X^1, X^2, ...\right]$, the Hicks-neutral efficiency level (A) and wet-bulb temperature (WBT) so that $Y = h\left[X(\text{WBT}), A(\text{WBT})\right]$. Assume that the firm or the household produces a homogeneous good with Cobb-Douglas technology:

$$Y_{jt} = A_{jt}(\text{WBT})\Pi_k(X^k_{jt}(\text{WBT}))^{\theta_k} \tag{S1}$$

Temperature WBT could affect marginal product of labor through its effects on TFP—which can be thought of as weighted average of capital productivity and labor productivity (Zhang et al. 2018), as well as on inputs via, for example, inducing worker absenteeism and reducing working hours (Somanathan et al. 2021; Graff Zivin and Neidell 2014).

To measure marginal product of labor, one can take natural logs of equation (S1) and obtain the empirical model:

$$y_{jt} = \theta_0 + \sum_k \theta_k x^k_{jt} + u_{jt} \tag{S2}$$

where y_{jt} is the log of value-added or gross revenue for firm or household j in year t, x^k_{jt} denote the log of k inputs. α_k is the output elasticity of the corresponding input k that need to be estimated. u_{jt} is the error terms. $ln(A_{jt}) = \alpha_0 + u_{jt}$ where $u_jt = \omega_{jt} + \eta_{jt}$. ω_{jt} is the household or firm productivity shock and the residual η_{jt} is assumed to have standard properties.

Estimating equation (S2) using Ordinary Least Squares (OLS) might be biased because of selection and simultaneity. Firms with lower productivity are more likely to exit the market, thus resulting in selection bias. In addition, firms can decide the levels of inputs based on their (partial) observation on productivity that is not observed by the econometrician.

To deal with these concerns, one can apply the approach proposed by Olley and Pakes (1996) (henceforth, OP) and Levinsohn and Petrin (2003) (henceforth, LP). The idea of OP approach is to use the survival rate of a firm to correct for selection bias and to use investment as a proxy for unobserved productivity shock to correct for simultaneity. This method assumes that investment (conditional on capital stock) is a strictly increasing function of the scalar, firm-level unobserved productivity shock, which means that one can invert the unconditional investment demand function and control for the unobserved productivity shock by conditioning on a non-parametric function of capital and investment. Similarly, LP approach assumes that intermediate goods are a strictly increasing function of a scalar, firm-level unobserved productivity shock. As discussed by Ackerberg, Caves, and Frazer (2015) (ACF), both OP and LP methods may suffer from functional dependence problems, that is, the condition underlying the first stage estimation may not identify the coefficients of variable inputs ("the colinearity problem"). The authors instead propose alternative procedure, which requires lagged values (e.g., lagged investment) for the estimation of the production function.

Equation (S2) can then be separately estimated for three groups: informal agriculture, informal non-agriculture, and formal non-agriculture. With the estimated input elasticity, one can derive the marginal product of labor for firm or household j in year t:

$$MP_{Ljt} = \hat{\theta}_l \frac{y_{jt}}{l_{jt}} \tag{S3}$$

and study the relationship between temperature and marginal product of labor by estimating the following equation:

$$MP_{Ljt} = f(\mathrm{WBT}_{pt}) + g(R_{pt}) + \gamma_p + \gamma_{rt} + \varepsilon_{prt} \tag{S4}$$

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015). "Identification properties of recent production function estimators". *Econometrica* 83.6, pp. 2411–2451.

Adhvaryu, A., N. Kala, and A. Nyshadham (2020). "The light and the heat: Productivity co-benefits of energy-saving technology". *Review of Economics and Statistics* 102.4, pp. 779–792.

Araar, A. and J.-Y. Duclos (2007). *DASP: Distributive analysis stata package*. Tech. rep. PEP, World Bank, UNDP and Université Laval.

Bolton, D. (1980). "The computation of equivalent potential temperature". *Monthly Weather Review* 108.7, pp. 1046–1053.

Dang, H. A. and P. W. Glewwe (2018). "Well Begun, But Aiming Higher: A Review of Vietnam's Education Trends in the Past 20 Years and Emerging Challenges". *Journal of Development Studies* 54.7, pp. 1171–1195.

GSO (n.d.[a]). *Vietnam (Household) Living Standards Surveys 1993-2018*. Data Set.

— (n.d.[b]). *Vietnam Annual Enterprise Census 2002-2016*. Data Set.

Levinsohn, J. and A. Petrin (2003). "Estimating production functions using inputs to control for unobservables". *Review of Economic Studies* 70.2, pp. 317–341.

LoPalo, M. (forthcoming). "Temperature, worker productivity, and adaptation: evidence from survey data production". *American Economic Journal: Applied Economics*.

McCaig, B. and N. Pavcnik (2021). "Entry and Exit of Informal Firms and Development". *IMF Economic Review* 69.3, pp. 540–575.

Minnesota Population Center (2015). *Integrated Public Use Microdata Series, International: Version 6.48 Vietnam Population and Housing Census 1989*. Data Set.

Nickell, S. (1981). "Biases in dynamic models with fixed effects". *Econometrica: Journal of the econometric society*, pp. 1417–1426.

Olley, G. S. and A. Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry". *Econometrica* 64.6, pp. 1263–1297.

Stull, R. (2011). "Wet-bulb temperature from relative humidity and air temperature". *Journal of Applied Meteorology and Climatology* 50.11, pp. 2267–2269.

Zhang, P., O. Deschenes, K. Meng, and J. Zhang (2018). "Temperature effects on productivity and factor reallocation: Evidence from a half million Chinese manufacturing plants". *Journal of Environmental Economics and Management* 88, pp. 1–17.