Climate Change and Intersectoral Labor Reallocation in a Developing Country

Trinh Pham*

December 9, 2023

Please click here for the latest version

Abstract

I present a simple model to reconcile different findings in the recent climate-employment literature, incorporating roles for trade openness and intersectoral switching costs. Using Vietnamese data, I provide new evidence consistent with the model's predictions. In most places, heat induces an outflow of workers from agriculture to non-agriculture in both the short and long terms. While all workers are equally likely to take an informal non-agricultural job, younger individuals incur lower costs and comprise most of those moving into formal non-agriculture. In places less integrated into global markets, the reallocation runs in the opposite direction due to general equilibrium effects.

Keywords: temperature, agriculture, informality, frictions, structural transformation, Vietnam

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

*Dyson School of Applied Economics and Management, Cornell University. Email: tp347@cornell.edu; 410 Warren Hall, 137 Reservoir Ave, Ithaca, NY 14853, United States. I am grateful to my advisors: Chris Barrett, John Hoddinott, and Ariel Ortiz-Bobea, for their guidance and support. For helpful comments and suggestions on previous versions of the paper, I also thank Jonathan Colmer, Teevrat Garg, Paul Glewwe, Maulik Jagnani, Brian McCaig, Arman Rezaee, Vis Taraz, Catherine Wolfram, as well as seminar participants at Cornell University, Midwest International Economic Development Conference (MWIEDC) 2023, Harvard Climate Economics Pipeline Workshop 2023, Agricultural & Applied Economics Association (AAEA) Annual Meeting 2023, LSE Environment Week 2023, and North East Universities Development Consortium (NEUDC) Conference 2023. Any remaining errors are mine.

1 Introduction

Developing countries are often characterized by a large informal sector (La Porta and Shleifer 2014; Ulyssea 2020) and frequent movement of workers into and out of marginal jobs, including self-employment, informal and low-earnings wage work (Donovan, Lu, and Schoellman 2023). These countries are also particularly affected the most by climate change, resulting in different patterns of labor reallocation across economic sectors. In some settings, changing climatic conditions disproportionately harm agricultural production and are associated with workers moving into other sectors in the short run (Colmer 2021; Liu, Shamdasani, and Taraz 2023), with some suggestive evidence that most of the reallocation is absorbed into small firms in the informal sector (Colmer 2021). In other settings, similar conditions lead to workers less engaging in non-agricultural activities in both the short (Jessoe, Manning, and Taylor 2018) and long runs (Liu, Shamdasani, and Taraz 2023).[1] Understanding to what extent climate induces labor reallocation and in which direction, and among which types of jobs is central to thinking about total climate change damages, as well as the future of jobs and food security in developing countries.

This paper provides a theoretical framework to reconcile the seemingly conflicting results in the recent literature, and generates new empirical evidence capturing heterogeneous effects of climatic conditions on structural transformation in the context of a developing country, Vietnam. A key feature of the empirical analysis is the utilization of both short-run weather variation and long-run climate variation to study the relationship between weather shocks and climate change, and labor allocation across economic sectors for different demographic groups and markets, thereby illustrating the role of market integration and labor market frictions in the climate–sectoral employment relationship.

The framework is developed as a simple general equilibrium model describing a two-sector economy, with heterogeneous individual productivity in non-agriculture, following Roy (1951). The model encompasses one of the main mechanisms suggested in the literature as a potential driver of structural transformation (Kongsamut, Rebelo, and Xie 2001) while taking into consideration the degree of openness of the economy (Matsuyama 1992). The model predicts that a larger loss in absolute value in agricultural productivity relative to non-agricultural productivity due to extreme

[1] Emerick (2018) finds asymmetric short-run effects of negative and positive rainfall shocks to local agricultural production in India. In particular, positive shocks lead to a decrease in agricultural work but negative shocks do not increase agricultural work.

temperatures can lead to a decrease in the agricultural labor share if the economy is open such that local supply shocks do not affect commodity prices. Furthermore, in this case, the rate of labor reallocation out of agriculture is decreasing in the costs associated with working in non-agriculture, which can be inferred from the observed earnings gains among workers who switched out of agriculture. In contrast, when the economy is sufficiently closed such that prices are endogenously determined by local supply and demand forces, a decrease in agricultural productivity can induce product price changes that lead to an increase in this sector's labor share. In sum, the model predicts that the direction and extent of climate change's impacts on sectoral labor allocation depend crucially on the economy's degree of openness and on the worker-specific costs of working in the non-agricultural sector.

To test the model's predictions, I rely on data from Vietnam, a lower-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.).[2] Since its economic reforms in the late 1980s, the country has been increasingly integrated into the global economy, becoming a leading rice exporter. The economy, however, is characterized by large variation in the level of market integration across provinces.[3] Likewise, the rate of change in sectoral employment shares differs substantially across age groups. The increase in formal non-agricultural employment share arises largely for younger birth cohorts entering the labor market more than for prior cohorts of workers at those same ages, even controlling for educational attainment. For informal non-agriculture, in contrast, the change in employment share is largely due to economy-wide trends, individuals of all birth cohorts move into this sector over time. Together, these features make Vietnam an ideal setting to explore the demographically and geographically heterogeneous effects of climate change on intersectoral labor reallocation.[4]

[2]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.

[3]Calculation using household surveys suggests that the correlation coefficient between the local price and the world price of rice ranges from 0.1 to 0.9 across provinces. See Appendix B3 for details.

[4]I do not distinguish between informal and formal agriculture because agricultural production in developing countries is predominantly carried out by smallholder farmers, who generally lack 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.

I assemble a province-age group panel dataset spanning nearly three decades from 1992 to 2018, based on 11 rounds of nationally and provincially representative household surveys, 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 supplement this dataset with other province-level panel datasets on migration, crop yields, cultivated land, as well as formal firm-level censuses. The use of sub-national datasets allows me to track local market responses, focusing on the allocation within a province, which I show to be an important, relevant empirical margin in this setting.[5] I proxy for trade openness of a province using the distance from its geographical centroid to the nearest (three) major seaports.[6] This measure is highly correlated with the price correlation between local and world rice–the country's main staple food, but is not associated with any difference in returns to non-agricultural work relative to agricultural work across spaces once individual selection into sectors of employment is controlled for.[7]

I document that provinces with a larger increase in extreme temperatures experienced a larger reduction in agricultural employment share. Such a correlation appears stronger for younger workers and in areas close to the country's major seaports. These observations are consistent with the theoretical prediction that temperature change can accelerate an outflow of labor from agriculture under certain conditions and that the rate of reallocation varies across groups who incur different costs of movement. However, there could be factors confounding this relationship. For example, the expansion of the education system and differential educational attainment across age groups over time might conflate the temperature effects.

To establish plausibly causal effects of temperature change on sectoral employment share by age group within Vietnamese provinces, I adopt two empirical approaches. The first approach exploits year-to-year variation in weather, while controlling for analysis unit fixed effects. This approach follows the recent climate 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

[5] In Section 4.4, I provide evidence that most of the structural change in Vietnam happens within provinces, and temperatures do not significantly affect inter-provincial migration rates.

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

[7] If anything, the return to informal non-agricultural employment relative to agricultural employment is higher in remote places than places close to major ports. There is no statistical difference in the relative return to formal non-agricultural employment across markets. More details in Appendix B3.

group cell. I call this the "panel approach." 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. I label this the "long differences approach." 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.

Two main findings emerge. First, temperature change significantly affects sectoral labor reallocation in both the short and long terms, and these effects happen entirely at the higher end of the temperature distribution.[8] While cold temperatures do not affect 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. Estimates using the panel approach show that every additional standard deviation increase in degree days above 27°C–the 97$^{\text{th}}$ percentile of the historical distribution–leads to a reduction of 0.05 points in provincial-level agricultural employment shares, and increases of 0.02 and 0.03 points in informal and formal non-agricultural employment shares, respectively. These amount to approximately 10-14% of the corresponding sample means. There is no evidence of a temperature effect on the share of inactive and unemployed individuals.[9] These effects of temperature on sectoral labor shares are similar under the long differences approach. The increased risks of extreme temperatures, as opposed to a general warming, leads to a reduction in the agricultural employment share and increases in both formal and informal non-agricultural employment shares. The estimated effects of hot temperature are significantly larger under the long differences approach than the panel approach.

Second, the average temperature effects mask significant heterogeneity by age group depending on the type of work climate change-induced intersectoral migrants

[8] In general, the effects of other weather variables are qualitatively similar to that of hot temperatures. 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 imprecisely estimated and less robust to alternative specifications and thus are omitted from the discussion.

[9] 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 involuntarily unemployed.

enter. While workers of all ages are equally likely to move out of agriculture and into informal non-agriculture in response to hot temperatures, younger workers comprise most of those who take formal non-agricultural jobs. In other words, most of the climate-induced labor reallocation takes place among agriculture and informal non-agriculture. These heterogeneous effects hold both in the short and long terms.

Supporting evidence suggests that these results are driven by relative sectoral productivity loss and non-uniform labor market frictions. Specifically, the positive (negative) effects of hot temperatures on non-agricultural (agricultural) employment shares are concentrated in areas that are reasonably open to trade, where commodity prices are largely unaffected by temperature shocks. In less connected areas, the reallocation effects run in the opposite direction. Hot temperatures increase the employment share in agriculture and decrease the labor share in non-agricultural sectors, as well as household (nonfood) consumption expenditures. Additional analyses show that hot temperatures, on average, significantly reduce hours worked and labor productivity in agriculture, but not in other sectors. Taken together, these findings are consistent with the model's predictions that a negative shock to agricultural productivity has differential effects on sectoral employment shares across spaces depending on the degree of market integration.

The heterogeneous temperature effects by age group and sector of in-migration appear consistent with the existence of non-uniform labor market frictions, particularly costs that vary across sectors and age groups. Following the model's prediction on the link between observed gains in earnings and costs of working in the non-agricultural sectors relative to agriculture, I use a sample of workers who changed sectors of employment to infer the cost of working in informal and formal non-agriculture, separately for different age groups. The estimated 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. Thus, given the change in relative labor productivity induced by hot temperatures, workers of different age groups have a 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 contributes to several strands of the literature. Most closely, it joins the growing body of empirical work linking climatic conditions and intersectoral labor reallocation in developing countries.[10] Most related are the work of Colmer (2021) and

[10] Examples include Emerick (2018), Jessoe, Manning, and Taylor (2018), Colmer (2021), and Liu,

Liu, Shamdasani, and Taraz (2023). On the one hand, Colmer (2021) documents an outflow of workers from agriculture to other sectors when commodity prices are not affected by short-term temperature change. Liu, Shamdasani, and Taraz (2023), on the other hand, find a reduction in the non-agricultural employment share in response to rising temperature in the long term, and this effect is concentrated among isolated areas. Building off their work, this paper provides a simple theoretical framework to understand the heterogeneous effects of climate change across tradable and non-tradable markets, while also considering the role of labor market frictions to provide new insights on the type of jobs climate-induced sectoral migrants take. Empirically, I utilize both short and long-run variations in weather and climate, coupling that with more comprehensive, micro-level datasets of the entire economy. These enable me to better understand heterogeneity in worker-specific responses to unexpected weather shocks and anthropogenic climate change, as well as the underlying mechanisms.

Similar to Liu, Shamdasani, and Taraz (2023), I document a larger temperature effect in magnitude when using the long differences approach relative to the panel approach.[11] However, there are differences when interpreting these results. On the one hand, in isolated places where hot temperatures are associated with a contraction of non-agricultural employment, the larger effect in magnitude in the long run is consistent with the intensification hypothesis wherein liquidity constraints limit individuals' ability to smooth consumption (Liu, Shamdasani, and Taraz 2023). On the other hand, in other well-connected places where hot temperatures decreases agricultural labor share, the larger effect over the longer time frame and in hotter places is consistent with workers making forward-looking decisions in the face of the trend

Shamdasani, and Taraz (2023). Given that temperatures have detrimental effects on crop production (Schlenker and Roberts 2009), a common feature of this literature is the focus on temperature-induced agricultural productivity shocks proxied by changes in crop yields as the main mechanism. Here I find that the effect of hot temperatures on yields of rice 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 at the intensive margin in response to hot temperatures. To the extent that increased input uses are induced by rising temperatures (e.g., Aragón, Oteiza, and Rud (2021) and Jagnani et al. (2021)), this implies a bigger direct effect of temperature on agricultural labor productivity had these adaptation practices not adopted. These findings suggest that the heat's impact on agricultural labor might transcend the commonly studied land mechanism wherein lower crop yields drive labor reallocation out of agriculture.

[11] One particular concern is that while the long differences approach is appropriate to study responses in the agricultural sector because other inputs and/or the production function (e.g., soil and land) has not changed much over time, the same might not be true for the non-agricultural sector. To the extent that other factors are evolving over time in a way that is correlated with differences in temperature, this will bias the effects on non-agricultural labor shares. While I cannot perfectly rule out this concern, the larger effects (in magnitude) in the long run in both well-integrated and isolated areas, where the latter did not enjoy much change in the non-agricultural sectors, do not seem to support this.

that global warming is disproportionately affecting the agricultural sector, and that non-agricultural sectors are becoming more attractive.

By explicitly examining the role of labor market frictions, this paper provides a potential explanation for why climate change might have lasting impacts on long-term economic growth (Dell, Jones, and Olken 2012), as well as highlights its inequality consequences for labor market outcomes. Due to the high cost of moving into the formal non-agricultural sector, most of the climate change-induced inter-sectoral labor reallocation takes place among (informal) agriculture and informal non-agriculture—the two low productivity and low skill-intensive sectors, consistent with the generally high labor market flows among marginal jobs documented across developing economies (Donovan, Lu, and Schoellman 2023). Recent work 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 proportion of climate change-induced workers move into the informal sector in the short and long terms might reinforce the 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).

Previous work on inequality consequence of climate change typically focuses on gender, race, or ethnicity dimensions (e.g., Maccini and Yang 2009; Park et al. 2020; Pham 2022), with a common underlying mechanism being the interaction between the direct effects of climate anomalies and preexisting gender bias in intrahousehold resource allocation, or differential access to coping and mitigation strategies due to socio-economic constraints. I show that temperature effects also vary across age cohorts and are likely driven by differential costs of switching sectors. The relatively low productivity and the lack of social welfare system for workers in the informal economy suggest that climate change may disproportionately affect the welfare of older populations.

Finally, the paper connects to the broader literature studying the role of agricultural growth in structural transformation. Some scholars argue that labor reallocation out of agriculture to manufacturing and service results from faster agricultural productivity growth that generates demand for non-agricultural goods.[12] Other scholars, however, argue that agricultural growth could lead to different patterns in the context of small open economies because of specialization according to comparative advan-

[12] See Kongsamut, Rebelo, and Xie (2001), Gollin, Parente, and Rogerson (2002), and Ngai and Pissarides (2007) for theoretical contributions. Bustos, Caprettini, and Ponticelli (2016) provides empirical evidence in the context of Brazil.

tage (Matsuyama 1992). The phenomenon studied here disproportionately impacts labor productivity in the agricultural sector, leading to different patterns of intersectoral labor reallocation across tradable and non-tradable markets. On the one hand, these results support the hypothesis that climate change widens the preexisting differences in intersectoral labor productivity, which could help drive structural transformation (Barrett, Ortiz-Bobea, and Pham 2023). But they also underline the model simulation-based findings of Nath (2020) that reducing trade barriers may be critical to mitigating climate impacts in low- and middle-income countries.

The remaining of the paper proceeds as follows. Section 2 presents the theoretical model. Section 3 describes the data and provides descriptive patterns of the temperature-employment relationship. Section 4 details empirical strategies and findings, reports results from a series of robustness checks and other analyses on migration, education, and gender. Section 5 explores potential mechanisms. Section 6 concludes.

2 A Simple Model of Frictional Labor Reallocation

In this section, I present a simple model of frictional labor reallocation. The model generates testable predictions for the temperature-sectoral employment share relationship in small open and closed economies, while also providing an interpretation of the observed gains in earnings when changing sectors of employment.

2.1 Environment

Endowment. The economy is composed of a unit measure of individuals who live for two periods. In each period, individuals are endowed with one unit of time and supply labor inelastically to one of the two sectors: agriculture (indexed g) and non-agriculture (indexed n).[13]

Working in non-agriculture is associated with a per-period cost that reduces non-agricultural earnings by a fraction τ in each period, where $0 < \tau < 1$. This cost can be interpreted as (i) taxes and social security contribution that are specific to (formal) non-agriculture, (ii) amenity cost, (iii) psycho-social costs associated with, for example, the exclusion from informal insurance networks (Munshi and Rosenzweig 2016), or the net of several of these effects.

[13] These assumptions are based on the fact that the share of non-employment, which includes inactive and involuntary unemployed individuals, remains relatively stable over the study period. Furthermore, as will be shown in the next section, this share is not significantly impacted by temperature changes.

Following Roy (1951), I assume that individuals are endowed with a bivariate skill vector $\{1, \varepsilon\}$ that represents the efficiency of their labor in agriculture and non-agriculture, respectively. In other words, individuals have identical productivity in agriculture and are only heterogeneous in the non-agricultural sector.[14] Furthermore, I assume that ε is distributed according to a Pareto distribution where the minimum of the support is normalized to the value one, $F(\varepsilon) = 1 - \varepsilon^{-\theta}$. The term $\theta > 0$ controls the variation in productivity with a lower θ implying more dispersion in individual productivity. This distribution assumption buys tractability of the model.

Production. Production in the agricultural and non-agricultural sectors uses labor services $L_{s\in\{g,n\}}$ as the sole-variable input.[15] The labor market is competitive. Productivity in agriculture and non-agriculture are Z_g and Z_n, respectively. The production functions in agriculture and non-agriculture are

$$\begin{aligned} Q_g &= Z_g G(L_g), \quad G(0) = 0, G' > 0, G'' < 0 \\ Q_n &= Z_n H(L_n), \quad H(0) = 0, H' > 0, H'' < 0 \end{aligned} \tag{1}$$

where $L_s = \int_{\Omega^s} \varepsilon_s dF(\varepsilon), \quad N_s = \int_{\Omega^s} dF(\varepsilon)$ denote the number of efficiency units and the total of workers employed, respectively, in sector s, with Ω^s representing the set of individuals employed in sector s.

Given the competitive labor market structure, firms' optimization implies that workers are paid the marginal product of their labor. Let the relative price of agricultural goods in equilibrium be p, then individual-level earnings in the two sectors are given by

$$\begin{aligned} y_g &= w_g = pZ_g G'(L_g) \\ y_n &= \varepsilon w_n = \varepsilon Z_n H'(L_n) \end{aligned} \tag{2}$$

where w_g and w_n represent the wages per efficiency unit in agriculture and non-agriculture, respectively.

[14]The assumption that non-agriculture is more human capital intensive than agriculture is consistent with evidence of sorting of high-skilled workers and larger returns to skills in this sector (e.g., Gollin, Lagakos, and Waugh (2014) and Herrendorf and Schoellman (2018)).

[15]In other words, land is assumed to be quasi-fixed. This assumption is consistent with evidence of non-changing cultivated land distribution in Vietnam over the study period (Liu et al. 2020). It also reflects the agricultural land protection policies in the country, in which the conversion of rice and agricultural land to other non-agricultural purposes is generally not allowed with very few exemptions that serve the national or public interests (e.g., Decree No.63/NQ-CP, Decision 124/QD-TTg, Resolution 17/2011/QH13).

Preferences. All consumers share identical preferences given by

$$u(c_g, c_n) = \alpha \log(c_g - \zeta) + (1 - \alpha) \log(c_n) \tag{3}$$

where c_g and c_n denote consumption of the agriculture good and the non-agricultural good, respectively. The parameter ζ represents the subsistence level of agricultural consumption and satisfies two conditions: (i) $Z_g > \zeta$ (the economy's agriculture is productive enough to provide at least a subsistence level of agricultural good to all consumers), and (ii) $\zeta > 0$ (non-homothetic preference assumption where income elasticity of demand for agricultural good is less than one).

Demand for the two goods by an individual satisfies $c_g = \zeta + \frac{\alpha}{(1-\alpha)p} c_n$. Aggregation over the unit mass of population yields

$$C_g = \zeta + \frac{\alpha}{(1 - \alpha)p} C_n \tag{4}$$

where the upper case letters denote aggregate consumption.

Workers. Workers are assumed to be initially randomly allocated to each sector and forced to work there in the first period. At the end of the first period, they make a switching decision to maximize their earnings in the second period. An individual with productivity vector $\{1, \varepsilon\}$ solves the problem

$$\max\{(1 - \tau)w_n \varepsilon, w_g\} \tag{5}$$

Definition of Equilibrium. A competitive equilibrium consists of sectoral choices and wages per efficiency unit $\{w_g, w_n\}$ such that:

- Firms maximize profits taking wages as given (2)
- Workers optimally choose sector of employment taking wages as given (5)
- Markets clear

Characteristics of Equilibrium. The workers' problem implies that an individual with productivity vector $\{1, \varepsilon\}$ works outside of agriculture if and only if $(1-\tau)w_n\varepsilon \geq w_g$. Thus there exists a single cutoff value $\hat{\varepsilon}$ that determines which individuals work in agriculture in the second period.

$$\hat{\varepsilon} = \frac{w_g}{(1 - \tau)w_n} = \frac{pZ_g G'(L_g)}{(1 - \tau)Z_n H'(L_n)} \tag{6}$$

The share of workers and efficiency unit of labor in agriculture and non-agriculture are determined as

$$
\begin{aligned}
L_g &= N_g = \int_1^{\hat{\varepsilon}} dF(\varepsilon) = 1 - \left[\frac{pZ_gG'(L_g)}{(1-\tau)Z_nH'(L_n)}\right]^{-\theta} \\
L_n &= \int_{\hat{\varepsilon}} \varepsilon dF(\varepsilon) = \frac{\theta}{\theta-1}\left[\frac{pZ_gG'(L_g)}{(1-\tau)Z_nH'(L_n)}\right]^{1-\theta} \\
N_n &= \left[\frac{pZ_gG'(L_g)}{(1-\tau)Z_nH'(L_n)}\right]^{-\theta}
\end{aligned}
\tag{7}
$$

2.2 Theoretical Predictions on the Temperature-Labor Share Relationship

In this section, I analyze the response of agricultural and non-agricultural employment in general and by age group to changes in extreme temperatures, generating empirically testable predictions. Extreme temperatures affect sectoral labor reallocation through their effects on sector-specific productivity Z_g and Z_n. While the adverse effect of heat on agricultural productivity is well-documented in the empirical literature (Schlenker and Roberts 2009), there is a dearth of evidence on the relationship between heat and non-agricultural productivity from developing countries, although existing evidence does suggest a negative heat impact on productivity of climate-exposed industries (Somanathan et al. 2021; LoPalo 2023).

Prediction 1: Small Open Economy. If the economy is sufficiently close to a small open economy in the absence of trade barriers and extreme temperatures disproportionately affect agricultural productivity, then

(a) extreme temperatures reduce the employment share of agriculture
(b) extreme temperatures increase the employment share of non-agriculture
(c) the reallocation effect (in magnitude) induced by extreme temperatures is decreasing in the cost of working in non-agriculture

Proofs are in Appendix A. Intuitively, when the economy is sufficiently close to a small open economy, the relative agricultural price is exogenously determined by the world market $\bar{p}$.[16] Changes in relative labor productivity loss caused by extreme temperatures induce workers to move away from the less productive sector to relatively more productive sectors. This happens differently by worker cohort and sector

[16] An implicit assumption is that local climate-related shocks are worse than elsewhere. If, on the other hand, climate-related shocks disproportionately affect agricultural production in other exporting economies, this might result in increased demand for the current economy's agricultural goods. In that case, the relative agricultural price would rise, leading to an increase in agricultural labor share.

of in-migration if labor market frictions vary among cohorts and sectors. For example, age-based discrimination practices in the hiring and firing processes among low-skilled labor-intensive employment in the formal non-agricultural sector might prevent the reallocation of older agricultural workers to this sector.[17]

Prediction 2: Closed Economy. If the economy is sufficiently closed and extreme temperatures disproportionately affect agricultural productivity, then under non-homothetic preference assumption

(a) extreme temperatures increase the employment share of agriculture
(b) extreme temperatures decrease the employment share of non-agriculture
(c) the reallocation effect (in magnitude) induced by extreme temperatures is decreasing in the cost of working in non-agriculture

Proofs are in Appendix A. When the economy is sufficiently closed, the relative agricultural price p is endogenously determined by local supply and demand. As agriculture constitutes a substantial share of the local economy and provides an important source of income for a large share of the population, negative shocks to agricultural productivity results in a reduction in income. When the income elasticity of demand for the agricultural good is less than one, there is a shift towards consumption of agricultural goods–the mechanism highlighted in previous literature as a potential driver of structural change (e.g., Kongsamut, Rebelo, and Xie 2001; Gollin, Parente, and Rogerson 2002). This channel also appears consistent with findings by Liu, Shamdasani, and Taraz (2023), who found the agricultural labor share increases in response to rising temperatures in India.[18]

[17] In Vietnam, some age-based discriminatory practices have been common, especially among low-skilled employment. On leading job sites such as vn.indeed.com, it is not difficult to find job advertisements with the keyword "lao động phổ thông" ("low-skilled workers" in Vietnamese) that restrict applications to a specific age range (e.g., 18-35, 18-40). Despite no comprehensive reports on age-based discrimination in labor and employment, existing studies conducted by the Ministry of Labor, Invalids and Social Affairs found that employers prefer recruiting young workers (See, for example, https://www.ilo.org/hanoi/Informationresources/Publicinformation/newsitems/WCMS_647834/lang–en/index.htm). According to a survey of 64 enterprises, the General Confederation of Labor documented that employers tend to terminate contracts with older workers to cut down wages and social security contribution, which is a function of their base salary and seniority pay (https://ldldsoctrang.soctrang.gov.vn/tin-hoat-ong/-/asset_publisher/s5406b90MKP4/content/van-e-sa-thai-lao-ong-tren-35-tuoi?).

[18] Alternatively, without non-homothetic preferences, a sub-unit elasticity of substitution across agricultural and non-agricultural goods may also lead to a shift of employment from non-agriculture to agriculture if the latter incurs a relatively larger productivity loss in magnitude caused by extreme temperatures. This mechanism is formalized by Ngai and Pissarides (2007).

2.3 Observed Gains among Switchers

The theoretical predictions involve estimating the cost of working in non-agriculture. In this section, I show that this cost can be inferred from the observed gains in labor earnings among the sample of workers who change sectors of employment.

For simplicity, let us assume that movement only happens from agriculture to non-agriculture, which is the dominant switching direction for individuals who do switch (Herrendorf and Schoellman 2018; Hamory et al. 2021).[19] Define the gains from switching out of agriculture for an individual worker as the difference in log earnings between agriculture and non-agriculture

$$\text{ind. gains} = \left[\log Z_n + \log H'(L_n) - \log Z_g - \log G'(L_g) - \log p\right] + [\log(\varepsilon) - \log(1)] \tag{8}$$

The first term is the benefit of working in the non-agricultural sector, relative to the agricultural sector. The second term reflects the relative productivity advantage that the individual enjoys in non-agriculture, reflecting selection on comparative advantage. Averaging across all switchers and applying equation (6) with properties of the Pareto distribution give us the average gains when moving from agriculture to non-agriculture[20]

$$\text{average gains among switchers} = \frac{1}{\theta} - \log(1-\tau) \tag{9}$$

Because the initial allocation of workers is random (and might thus be not optimal), the average gains in earnings also capture the composition of individuals who switch out of agriculture in equilibrium (selection according to comparative advantage). If, instead, workers are assumed to self-select in the first period and only switch at the margin in the second period because of a change in relative sectoral produc-

[19]When the analysis is restricted to this specific direction, one can further introduce a fixed cost that individuals incur when a change in sector of employment (agriculture to non-agriculture) takes place. The gains in earnings then reflect two types of costs: a per-period cost of working in non-agriculture and a one-off switching cost to this sector. All the qualitative conclusions remain the same.

[20]In particular, the log transformation of a Pareto distribution is distributed exponentially with parameter θ. The average gains among switchers are

$$\begin{aligned}\text{average gains} &= \left[\log Z_n + \log H'(L_n) - \log Z_g - \log G'(L_g) - \log p\right] + \int_{\hat{\varepsilon}} \log(\varepsilon) f(\varepsilon|\mathcal{E} \geq \hat{\varepsilon}) d\varepsilon \\ &= \left[\log Z_n + \log H'(L_n) - \log Z_g - \log G'(L_g) - \log p + \log(\hat{\varepsilon})\right] + \int_{\hat{\varepsilon}} \log\left(\frac{\varepsilon}{\hat{\varepsilon}}\right) f(\varepsilon|\mathcal{E} \geq \hat{\varepsilon}) d\varepsilon \\ &= -\log(1-\tau) + \frac{1}{\theta}\end{aligned}$$

tivity (e.g., induced by extreme temperatures), then any individual who is observed to move must have productivity approximately equal to the threshold $\hat{\varepsilon}$.[21] In other words, the marginal switcher assumption further implies that

$$\text{average gains among switchers} \approx -\log(1-\tau) \tag{10}$$

Taken together, under these assumptions, equations (9) and (10) both suggest that if the observed gains from working in non-agriculture relative to agriculture among workers who changed employment sectors are small, then the costs must be small and vice versa.

3 Data and Descriptive Patterns of Temperature and Employment

3.1 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 B. The main sources of data include the Vietnamese household surveys, the population and formal firm census, and the global climate and weather reanalysis ERA5 database.

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.[22] 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.[23] The analysis sample includes workers with

[21] Lagakos et al. (2020) and Schoellman (2020) discuss these points in the rural-urban migration setting.

[22] The earliest household surveys in 1990s are not representative at the province level, I re-visit this point in the robustness checks section. 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%.

[23] 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

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.[24]

The key variables of interest are the sector in which an individual was working during the reference period and their working hours. These variables are 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.[25] 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 (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.[26]

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 (intensive margin). On the employer side, informality implies that firms do not register with the government (extensive margin). A cross-check whenever possible suggests that the notion of informality employed in

Schoellman (2023) construct a dataset of gross labor market flows of individuals of up to 6-9 months for a sample of 45 countries.

[24]With this restriction, this paper likely overlooks a potentially important margin: the school-to-work transition of young individuals.

[25]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.

[26]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 (2018). Employment in cooperatives and collectives since 2000 contributes to less than 1% of the total employment of adults 24-64 years old.

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 construction.[27] If climate change induces reallocation of workers into temporary jobs, as what Colmer (2021) found in the case of short-run increase in average temperature, 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.

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 $0.25° \times 0.25°$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).[28]

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 boundaries in Vietnam over the study period, and most of the changes happens in the case of splitting, I use the

[27] See Appendix B2 for a breakdown of the share of informal workers in formal firms by industry. 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)

[28] 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)).

original administrative units in 1993, which gives a consistent sample of 52 provinces over the study period.[29]

Other Data. In addition to employment and weather data, I use the Enterprise Census, which covers all formal firms in Vietnam during the period 2002-2016 to construct a firm-level and a province-level longitudinal dataset of labor productivity, as proxied by revenue per worker for the formal non-agricultural sector. Finally, I compile a dataset of province-level migration rates, agricultural cultivated area and crop yields from statistical yearbook.

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.[30] 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.

3.2 Descriptive Patterns of Temperature and Sectoral Labor Share

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. At the same level of heat, a place with dry air has a lower wet-bulb temperature and feels cooler compared to a place with humid air, because the former

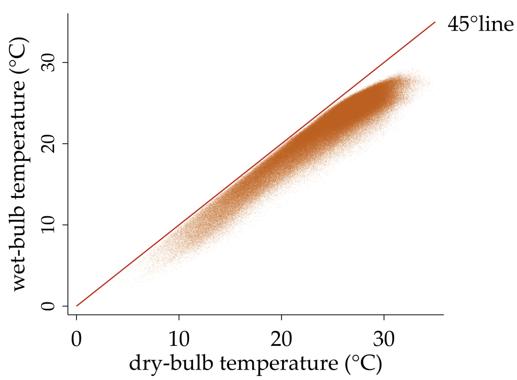
[29] An exception is the then Ha Tay province, which was merged into Hanoi city in 2008, and therefore I use the boundary of the new Ha Noi for consistency.

[30] 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.

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 wet-bulb temperature 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).

Vietnam is characterized by different climatic conditions that vary greatly between regions due to its diverse topography, long latitude, and influences from the East Sea. According to Beck et al. (2018), the country can be classified into seven climatic regions including tropical-rain forest, 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.[31] At the same level of dry-bulb temperature, there is significant variation in the level of humidity across provinces, which leads to a wide range of wet-bulb temperatures. For example, when air temperature falls between 27-30 °C, WBT ranges from a cool 18°C to a hot 28°C, which potentially have very different effects on human health and productivity (Figure 1). In what follows, I rely on the measure of wet-bulb temperatures and refer to it as temperature, unless otherwise stated.

Figure 1: Dry-bulb and Wet-bulb Temperatures (°C), 1980-2020



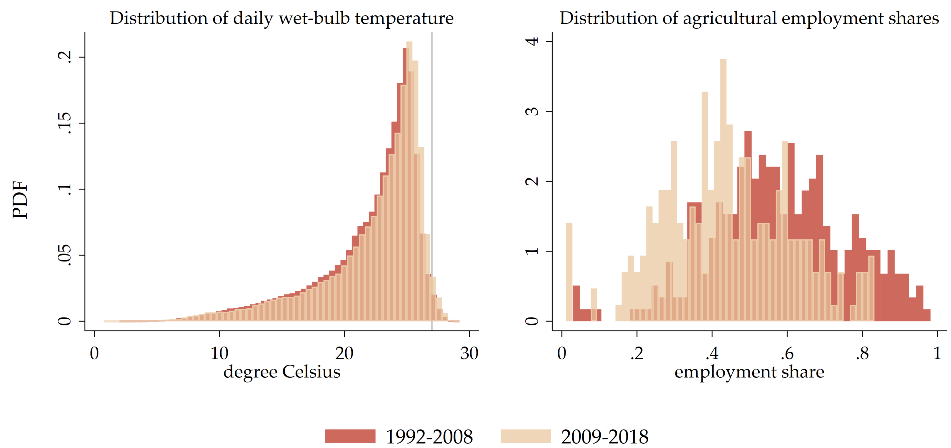
Notes: Each point represents the mean temperatures of a day of a province during 1980-2020.

Figure 2 shows a shift in the province-level temperature distribution and agricultural employment shares over time. The left panel documents temperatures not only

[31] There are 5-13 provinces per region.

increased in mean level over time, but also became more variable with more extreme hot and cold days. The right panel shows a steady decrease in agricultural labor shares across provinces.

Figure 2: Distribution of Wet-bulb Temperatures and Agricultural Labor Share across Provinces



I capture the change in temperature distribution over time using a statistical measure: Kullback-Leibler divergence (KLD, henceforth). A KLD value of zero implies two distributions are identical, and a greater value implies more difference between the two distributions.[32] I then decompose the overall difference between two distributions into the location and shape components. 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.[33]

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

[33]Appendix Figure D1 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 recent distribution is similar to the reference distribution in Province A. In Province B, although less

Distinguishing location and shape differences is likely important to assess the effect of temperature 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).

Panel A of Figure 3 presents the change in temperature distribution in Vietnamese provinces over two periods: 1992-2008 and 2009-2018 using the measure of shape difference. Temperature change 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 and relatively larger increase in living standards, as proxied by household consumption, between the two periods. These relationships appear stronger in areas close to major seaports. Panel B (left) confirms a clear negative association between the shape difference measure and the change in agricultural employment share at the province level. The right panel likewise plots the negative relationship between a more conventional measure of extreme temperatures (degree days above 27°C) and sectoral employment share change.

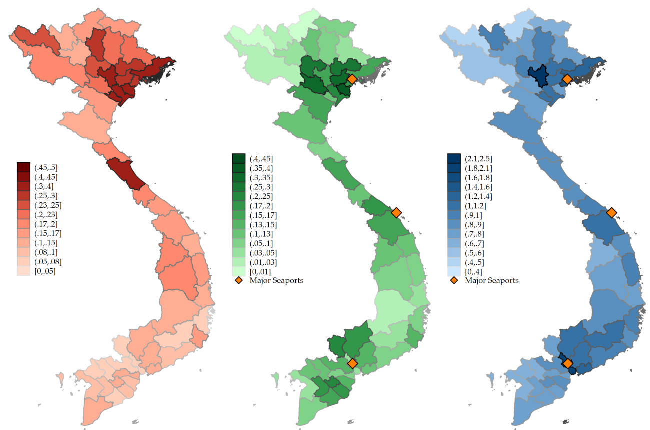
A key feature of the evolution of sectoral employment shares in Vietnam is the stark difference in the rate of labor reallocation across age cohorts. Panel A of Figure 4 shows the share of workers in each sector for four 4-year intervals from 1989 to 2018 and for five 4-year age intervals. As seen, those 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%. 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.

Figure 4 Panel B plots the relationship between changes in temperature distribution and changes in sectoral employment share for three age groups: 24-39, 40-54,

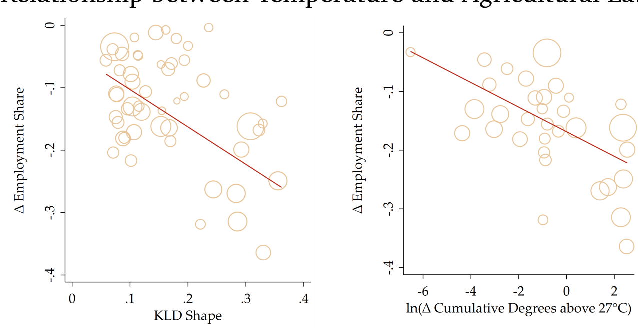
precisely estimated, there appears to be a significant change in the shape of the distribution with fewer mid-range days and and more days on the right tail.

Figure 3: Change in Temperature Distribution and Labor Shares 1992-2018

Panel A: Shape Difference (Left), Decrease in Agricultural Labor Share (Middle), and Increase in Household Consumption (Right)



Panel B: Relationship between Temperature and Agricultural Labor Share



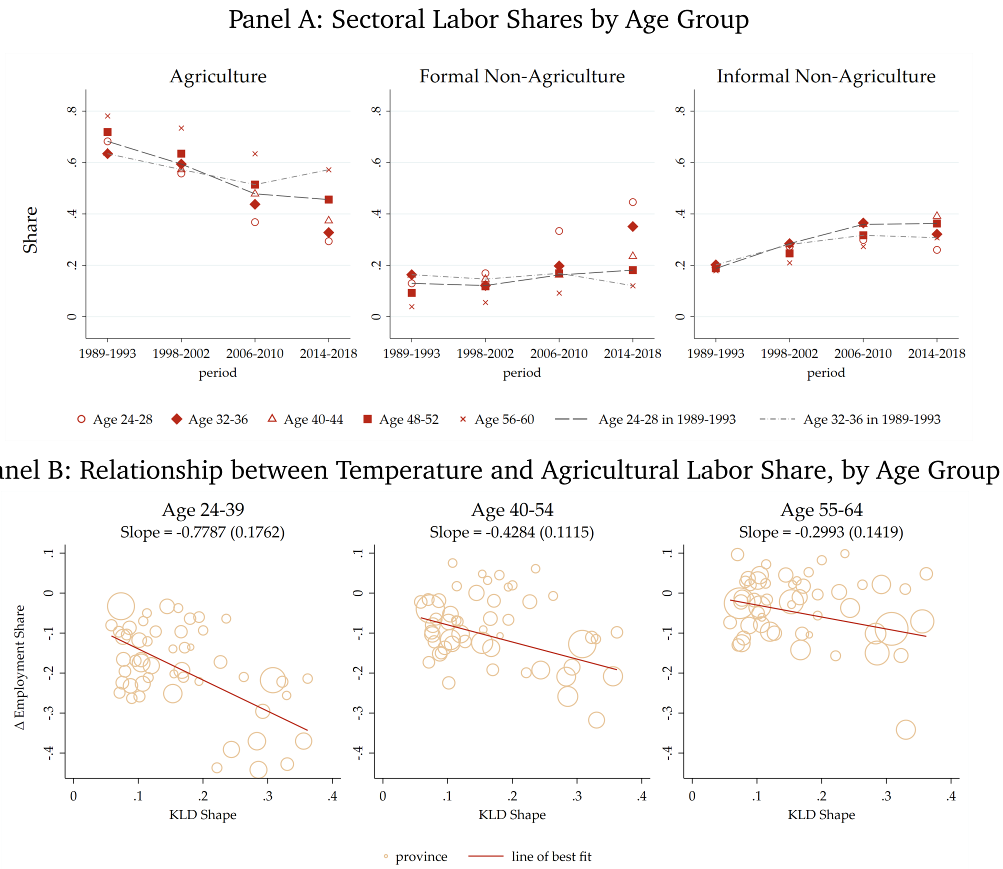
Notes: In panel A, darker color denotes larger magnitude. Panel B shows the line of best fit from the regression of change in sectoral employment shares on change in temperature distribution, as proxied by shape difference (left) and change in extreme temperatures (right), with each circle representing a province. The size of the circle is proportional to sampling weights.

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 labor share is stronger for the younger group.

4 Empirical Analysis

In what follows, I provide empirical evidence on the plausibly causal effects of temperature change on structural transformation in Vietnam. The basic correlation in the data reported in the previous section implies that provinces where there were

Figure 4: Change in Temperature Distribution and Labor Shares by Age Group



Notes: Panel A plots the share of workers in each sector for four 4-year intervals from 1989 to 2018 and for five 4-year age intervals. Change in agricultural and formal employment share largely follows younger cohorts entering the labor market more into this sector. 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. Panel B shows the line of best fit from the regression of change in sectoral employment shares on change in wet-bulb temperature distribution, as proxied by shape difference across age groups, with each circle representing a province. The size of the circle is proportional to sampling weights.

larger change in the temperature distribution (particularly extreme temperatures) experienced a reduction in agricultural labor share while non-agricultural employment expanded, and such a correlation is prevalent for areas close to the major seaports, as well as for younger workers. These findings are consistent with those predicted by the model. However, these correlations are not informative about the direction of causality. For example, changes in demographic characteristics, including educational attainment, that are correlated over time with rising temperatures might conflate the temperature effects.[34]

4.1 Empirical Strategy

To examine the causal effect of temperature change on sectoral labor allocation and other outcomes, I employ two different approaches, each of which relies on a different identification strategy and source of variation in the weather distribution.

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:

$$y_{aprt} = f(a, \mathrm{WBT}_{pt}) + g(a, \mathrm{R}_{pt}) + \gamma_{ap} + \gamma_{art} + \epsilon_{apt} \tag{11}$$

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.[35]

Equation (11) 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}. Across all estimations with equation (11), I weight the results by age group-specific population so that the coefficients correspond to an average person in the relevant age category.[36] I cluster standard errors at the province level to allow for potential serial correlation over time within each province (52 provinces). I also report Conley standard errors that allow for spatial correlation up to 150 km and serial correlation up to five lags in the error term

[34] Porzio, Rossi, and Santangelo (2022) 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.

[35] 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.

[36] 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 year t.

(Conley 1999).[37]

The focus of equation (11) 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:

$$f(a, \text{WBT}_{pt}) = \begin{cases} \sum_a \sum_{d=1}^{365} \beta_{a9}(9 - \text{WBT}_{dpt})\mathbb{I}_a & \text{if} \quad 0 \leq \text{WBT} < 9 \\ 0 & \text{if} \quad 9 \leq \text{WBT} < 27 \\ \sum_a \sum_{d=1}^{365} \beta_{a27}(\text{WBT}_{dpt} - 27)\mathbb{I}_a & \text{if} \quad \text{WBT} \geq 27 \end{cases} \tag{12}$$

With this function, β_{a9} and β_{a27} can be interpreted as the effect of one additional degree day below 9°C (DD9) or above 27°C (DD27), respectively, on sectoral employment shares of age group a over the 12-month reference period.[38] The 9°C-27°C range captures the middle 95% of the daily wet-bulb temperature distribution in the sample during the study period (Appendix Figure D2). The use of this function captures an agreement in the ergonomic literature that human performance loss from temperature is non-linear, with little or no loss associated with temperature increases in moderate temperature regimes and large loss associated with temperature increases in high temperature regimes. For example, findings from three meta analyses that human performance drops significantly once wet-bulb globe temperature is above 27°C (Hsiang 2010).

I also estimate a model where $f(.)$ is represented by cumulative temperature bins, degree day bins, and fourth order polynomials. These models provide sufficient flexibility to capture important non-linearity, as well as being relatively parsimonious with low demand on the data. The results from these alternative functional forms of temperature are similar to the baseline results.[39]

Identification Assumption. The validity of estimates based on equation (11) relies on the assumption that $\mathbb{E}\left[f\left(a, \text{WBT}_{pt}\right) \epsilon_{aprt} | g(a, \text{R}_{pt}), \gamma_{ap}, \gamma_{art}\right] = 0$. 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 extents.

[37] 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).

[38] 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).

[39] See Appendix C1 for the construction of these measures and the set of results.

The inclusion of these fixed effects are important for the following reasons. First, it addresses the concern about age group differentiated differences in educational attainment, in particular, 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. Second, it controls 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 that aims to boost industrial sector, generating 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. It also addresses the case in which 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.[40] The identifying variation is assumed to be orthogonal to unobserved determinants of sectoral employment in each age group-province cell.[41]

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:

$$\Delta y_{apr} = f_{\mathrm{LD}}(a, \Delta \mathrm{WBT}_p) + g_{\mathrm{LD}}(a, \Delta \mathrm{R}_p) + \gamma_{ar} + \epsilon_{apr} \tag{13}$$

where Δy_{apr} represents the change in sectoral employment shares of age group a of province p between two sub-periods. The shares in each period are calculated as the average of the shares in each survey waves during that period.[42] 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 precipitation and its squared, as

[40] Examples of non-monetary value include flexible work schedules, paid leave, vacation.

[41] A concern with the use of multiple fixed effects is that they absorb a significant amount of weather variance. Appendix Table D1 shows that even with these set of fixed effects, the remaining variation in weather variables is substantial. In addition, as shown in Appendix C1, the following results are robust to alternative specifications in which different set of fixed effects is employed. In Section 4.3, I also explore the robustness of the results to inclusion of time-varying demographic characteristics, including educational attainment.

[42] For example, to compare 1992-2008 and 2009-2018, the first sub-period comprises of six waves of household surveys 1992/1993, 1998, 2002, 2004, 2006, and 2008, whereas the second sub-period comprises of five waves of household surveys 2010, 2012, 2014, 2016, 2018.

well as number of days with high wind speeds. Equation (13) 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 150 km.[43]

I estimate two specifications of equation (13). The first specification uses the measures of KLD location and shape differences, which are calculated using the full 12-month temperature distribution before the average interview timing during each sub-period (e.g., 1992-2008 and 2009-2018), to proxy for ΔWBT_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 province-specific reference year instead of individual-specific 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 average effects of province-level temperature change on province-level sectoral employment shares. As seen in Appendix Figure D3, in most provinces, the average interview was completed between July-August, which supports this interpretation.

Although KLD measure is useful to proxy for change in the temperature distribution over an extended period of time, they are not without drawbacks: the measure is difficult to interpret, and is not a metric.[44] To facilitate a comparison between the coefficients estimated using long-term change in climate and those estimated using short-term weather variation, I estimate equation (13) with the temperature response function being defined as:

$$f_{\text{LD}}(a, \Delta\text{WBT}_p) = \begin{cases} \sum_a \beta_{a9,\text{LD}} \Delta(9 - \text{WBT}_{pt}) \mathbb{I}_a & \text{if} \quad 0 \leq \text{WBT} < 9 \\ 0 & \text{if} \quad 9 \leq \text{WBT} < 27 \\ \sum_a \beta_{a27,\text{LD}} \Delta(\text{WBT}_{pt} - 27) \mathbb{I}_a & \text{if} \quad \text{WBT} \geq 27 \end{cases} \tag{14}$$

[43] I do not find any evidence of spatial correlation in residuals from the Moran test conducted after estimating equation (13) 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.

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

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 9°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.

Identification Assumption. The identification comes from within-region variation in changes in temperature and weather extreme between the two periods, which removes the effect of any time-invariant omitted variables at the province-age group level while also eliminating concerns over time-trending unobservables at the region level. Conditional on this assumption, the coefficient on temperature variables captures the causal effect of long-term change in the temperature distribution on sectoral employment allocation.

4.2 Empirical Results

4.2.1 Panel Approach

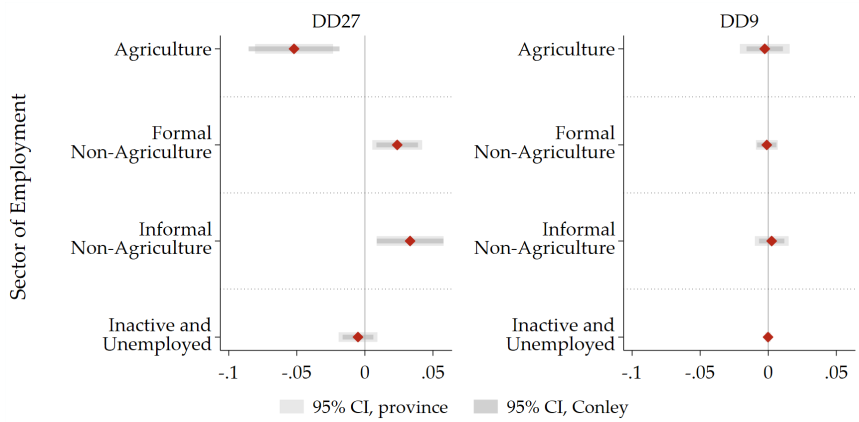
Average Temperature Effects. Figure 5 presents the results from estimating a variant of equation (11) in which the response functions $f(.)$ and $g(.)$ are not specific to any age group. As seen, cold temperatures do not affect sectoral labor shares. One standard deviation increase in degree days above 27°C, however, decreases the provincial-level employment share in agriculture by roughly 0.05 points (p-value $<$ 0.01), which amounts to approximately 10% of the outcome mean. The corresponding effects on formal and informal non-agricultural labor shares are 0.02 and 0.03 points, respectively (roughly 11-14% of the corresponding outcome means). 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. Because there is no significant change in the share of inactive and unemployed workers, these findings do not reflect a labor supply effect, but a labor reallocation effect.

Appendix Table C1 presents the results from the full model, suggesting a non-linear impact of precipitation on labor allocation. The impacts are of expected signs and smaller magnitude compared to hot temperatures, with increasing precipitation being associated with a decrease in the agricultural labor share and increases in the non-agricultural labor shares. These precipitation effects, however, are less precisely estimated. In what follows, I omit weather variables other than temperatures from the discussion.

To have a better understanding of which industry climate-induced migrants move into, I classify non-agricultural employment into three main groups: less, medium, and high skill-intensive industries.[45] This further breakdown of the destination sector suggests that roughly 75% of the labor reallocation induced by hot temperatures is into less skill-intensive industries (Appendix Table D2), including low-tech manufacturing (e.g., manufacture of food products and beverage, textile and apparel), less knowledge-intensive services (e.g., wholesale and retail, transport, accommodation and food services), construction and mining.

Figure 5: Wet-bulb Temperature and Sectoral Labor Share

Results from Panel Approach with Degree Days



Notes: This figure presents the effects (as point change) of 1 SD increase in degree days with wet-bulb temperature above 27°C (DD27), or below 9°C (DD9) on sectoral labor shares. Estimates obtained from a variant of equation (11) in which the response functions $f(.)$ and $g(.)$ are not specific to any age group. Unit of analysis is province-age group-year. All regressions 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 and Conley standard errors that allow for spatial correlation up to 150 km and serial correlation up to five lags are reported. Province distances are computed from province geographic centroids. All regressions use sampling weights. All regressions include sampling weights.

In Appendix Figures C5 and C6, I present additional results with the temperature functions being represented by cumulative temperature bins, degree day bins, and fourth-order polynomials. Across these alternative functional forms, the results are

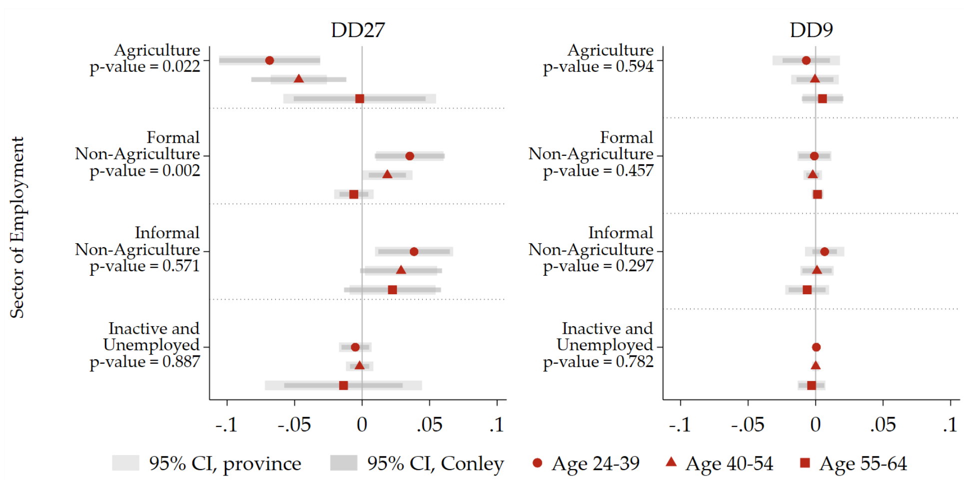
[45] Industries are ranked following the Statistical Classification of Economic Activities in the European Community. Less skill-intensive industries include those classified as low-tech manufacturing, less knowledge-intensive services, construction and mining; medium skill-intensive industries include those classified as medium-tech manufacturing and public utilities; high skill-intensive industries include those classified as high-tech manufacturing and knowledge-intensive services according to the Statistical Classification of Economic Activities in the European Community. The mean educational attainment of workers in each industry is presented in Appendix Table B3. For details, see Annex 3 – High-tech Aggregation by Statistical Classification of Economic Activities in the European Community (NACE Rev.2)

similar to those obtained from the parsimonious baseline model: all of the labor reallocation effects are driven by the higher end of the temperature distribution–the level above approximately 27°C.

Heterogeneous Effects by Age Group. Figure 6 shows the effects of temperatures on three group of workers, which are estimated from equation (11). Cold temperatures do not affect sectoral allocation of any group. Younger workers are less likely to work in agriculture in response to hot temperatures: one standard deviation increase in degree days above 27°C wet-bulb temperature decreases agricultural labor share for workers age 24-39 by roughly 0.06 points. The corresponding effect for workers age 40-54 is 0.04 points. The hot temperature impact on older workers is less precisely estimated, with the near-zero point estimate suggesting that they are virtually not affected. Correspondingly, the two younger groups also experience significant increases in the formal non-agricultural employment shares, with 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.

Figure 6: Wet-bulb Temperature and Sectoral Labor Share by Age Group

Results from Panel Approach with Degree Days



Notes: This figure presents the effects (as point change) of 1 SD increase in degree days with wet-bulb temperature above 27°C (DD27), or below 9°C (DD9) on sectoral labor shares of different age groups. Estimates obtained from equation (11). p-values from the F-test of significant age cohort differences using standard errors clustered at the province level are reported. The results are qualitatively similar when using Conley standard errors.

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 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 effect on non-employment for any group, which suggests that this is not an income channel effect, else labor supply should rise.

4.2.2 Long Differences Approach

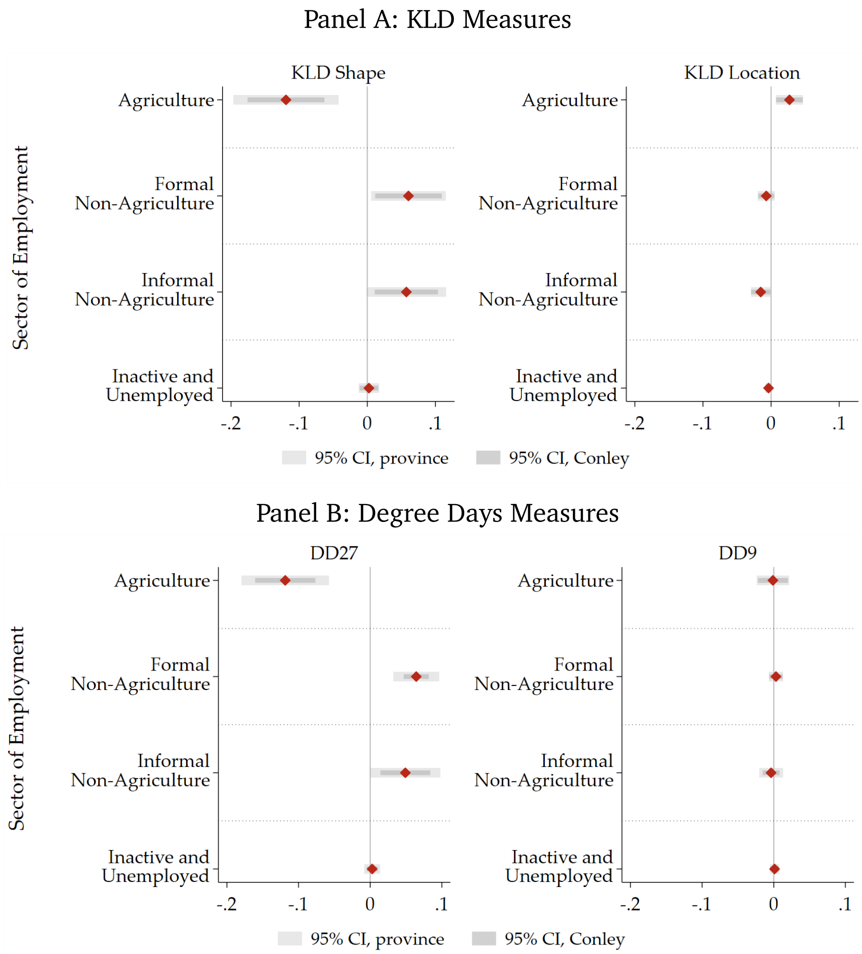
Figure 7 presents the temperature effects on the four key outcomes estimated using the long differences approach. Consistent with the panel results, controlling for other weather variables and region-specific age group-specific trends, provinces experiencing a larger change in the shape of temperature distribution between any two sub-periods see a larger reduction in the agricultural labor share and increases in the formal non-agricultural employment share, and these effects are all statistically significant at the 5% level when standard errors are clustered at the province level, and account for spatial and temporal correlation. The effects on the informal labor share are positive but less precisely estimated. No temperature effect on the share of inactive and unemployed workers is detected.

Using the shape and location components in the long differences specification allows one to examine the relative importance of "general warming" versus "increased risk of extreme temperatures" in inducing labor allocation. Evaluated at the sample mean, while the effects of shape difference are significant, general warming as proxied by location shifts have much smaller effects.[46] 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).

Figure 7 Panel B presents the long differences estimates with changes in degree days during the reference period being used as the key independent variables of interest, using the periods 1992-2008 and 2009-2018. Consistent with findings from the panel approach, cold temperatures do not have any significant effect on sectoral labor shares. Across agriculture and non-agriculture specifications, the long differences estimates of hot temperature effects are of larger magnitude than the respective panel estimates (p-values < 0.05). Evaluated at the sample mean, the point estimate of

[46] Furthermore, the statistically significant effects of location shifts are not robust to additional tests that will be discussed in the next section (See Appendix Figure C4).

Figure 7: Wet-bulb Temperature and Sectoral Labor Share: Long Differences



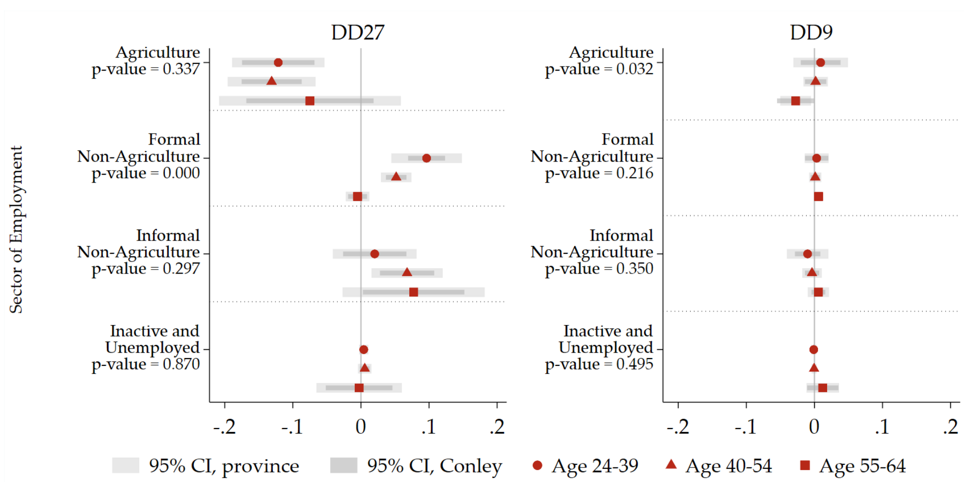
Notes: This figure presents the effects (as point change) of moving from a value of zero to the sample mean in KLD shape and location (Panel A) and the effects of 1 SD increase in degree days above 27°C or below 9°C (Panel B) on sectoral employment shares, which are obtained from estimating equation (13). Panel A and Panel B include the two sub-periods 1992-2008 and 2009-2018. Province distances are computed from province geographic centroids. All regressions include average sampling weights.

the hot temperature effects on agricultural labor share is -0.028 (95% CI = [-0.044, -0.016]) with the panel approach, and -0.074 [-0.112, -0.036] with the long differences approach. The corresponding effects on formal non-agricultural labor share are 0.013 [0.003, 0.023] (panel approach), and 0.041 [0.020, 0.060] (long differences approach). The hot temperature effects on informal non-agricultural employment share using the panel and long differences approach are 0.018 [0.005, 0.031] and 0.031 [0.0002, 0.061], respectively.

Similar to the panel approach, the long differences estimation also yields significant heterogeneous effects of hot temperatures on formal non-agricultural labor shares across age groups (Figure 8). Although I cannot reject the null of no differences in temperature effects for the agricultural labor share across age groups, the trend in the point estimates suggest that older workers are less responsive to hot temperatures. There is no evidence of differential temperature effects across age groups for informal non-agriculture, as well as no labor supply effect. There is little evidence of (differential) cold temperature effects on sectoral employment shares across age groups (Figure C12).

Figure 8: Wet-bulb Temperature and Sectoral Labor Share by Age Group

Results from Models with Degree Days, Long Differences Approach



Notes: This figure presents the effects (as point change) of 1 SD increase in degree days with wet-bulb temperature 27°C (DD27) on sectoral labor shares of different age groups. Estimates obtained from equation (11). p-values from the F-test of significant age cohort differences using standard errors clustered at the province level are reported. The results are qualitatively similar when using Conley standard errors.

4.3 Robustness Checks and Placebo Test

Robustness Checks. I assess the robustness of the main estimates on average and age-heterogeneous temperature effects on sectoral labor allocation to a number of deviations from the baseline specification in Appendix C1. First, the estimates are stable when I control for time-varying demographic characteristics that might influence sectoral employment, including share of male workers, share of ethnic minority workers, and educational attainment. Second, I obtain similar estimates with different sample restrictions, including removing observations from earlier household survey rounds which are arguably not representative at the province level. Similarly, the results are not sensitive to excluding observations of a specific province from the estimation. Third, the estimates are not driven by potential omitted variable bias caused by the high correlation in sectoral employment shares at the province level over time. Fourth, the results are robust to alternative 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 boundary) and alternative weather exposure definition. Fifth, the results from the panel approach are robust to different functional forms of temperature (e.g., fourth-order polynomials, cumulative temperature bins, and degree day bins), while the results from the long differences approach are robust to different period definitions.

Placebo Test. In Appendix C2, I verify using Monte Carlo-based permutation analyses that (11) and (13) provide correct inference and unbiased estimates of the temperature effects given the properties of the current data.

Additional Results using Dry-bulb Temperatures. In Appendix C3, I show that similar to the results using wet-bulb temperatures, hot dry-bulb temperatures are associated with a decrease in agricultural labor share and increases in formal and informal non-agricultural employment shares, but do not affect the share of unemployed and inactive workers. These coefficients, however, are less precisely estimated compared to when using wet-bulb temperatures.

4.4 Migration, and Effects by Education and Gender

Migration Analysis. In studying intersectoral labor reallocation, I have implicitly assumed that local markets are bounded at the province level. Previous research, however, has demonstrated the prevalence of human migration across spaces in response to climate change.[47] 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 discussed in Section 5.

I explore this concern by first conducting a decomposition exercise following McCaig and Pavcnik (2018). Results in Appendix Table D3 suggest that most of the structural change happens within provinces through this decomposition. In addition, while I do not have micro-level migration data for each age group over the study

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

period, I provide supporting evidence by estimating a variant of equation (11) using aggregate data from statistical yearbook, where the outcomes being the rates of migration, including out-migration, in-migration, and net-migration at the province level. Appendix Table D4 shows little evidence of cold and hot temperature effects on migration rates.

The findings based on the decomposition and migration analysis exercises suggest that inter-provincial migration is not a first-order response and provincial-level labor markets are relatively bounded. As a result, 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 (e.g., Maccini and Yang (2009) and Björkman-Nyqvist (2013)). In this context, however, I find no evidence of heterogeneous temperature effects on intersectoral labor reallocation by gender (Appendix Table D5). Likewise, I find limited evidence of differential temperature effects by education level, which again suggests that the heterogeneous results by age cohorts in the formal non-agricultural sector cannot be explained by differences in educational attainment across these groups.[48]

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 Hot Temperatures Accelerate Labor Reallocation out of Agriculture

According to the theoretical model, the results on average effects where hot temperatures induce reallocation of workers from agriculture to non-agricultural sectors are consistent with being predominantly driven by the relative labor productivity loss mechanism when the price is hold fixed. In what follows, I offer additional evidence to support this channel.

[48] I observe differential temperature effects on labor shares in informal non-agriculture. In particular, workers with a high school diploma are much less likely than peers without one to get an informal non-agricultural job in response to hot temperatures, which might reflect preference differences among individuals with different educational attainments.

The Labor Reallocation Effects Out of Agriculture Are Concentrated in Areas that Are More Integrated Into the World Market and Whose Prices Are Less Affected by Temperature Change. To begin, I test whether temperatures affect prices of rice–the country's main staple crop. I construct a province-level panel dataset of rice prices using the household surveys and I estimate equation (11) with the outcomes being mean and median price of rice at harvest, as well as mean and median price of rice sold by households in each province. Results in Appendix Table D6 shows that rice prices across the country are not significantly affected by temperature changes.

Next, 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 (11) and (13) where weather variables, including temperature, are interacted with an "open" indicator. I proxy for trade openness using two measures. The first measure is the distance from a province geographic centroid to the nearest major seaport. The second measure is the correlation coefficient between local agricultural price series, specifically rice price, and that of the world market.[49] The idea is that areas closer to major seaports and/or more integrated to the global economy are less affected by temperature-induced price and local demand effects, and thus most labor reallocation is driven by relative labor productivity loss.

Table 1 presents results from the panel and long difference approaches using distance to the nearest major seaport as a proxy for trade openness. As seen, the temperature effects on agricultural and non-agricultural employment shares are entirely driven by areas that are relatively more open to trade. 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 opposite effects are imprecisely estimated in the panel approach, they are intensified and become statistically significant at the 5% level when evaluated over a longer time frame. The point estimates of the temperature effects on informal and formal non-agricultural sectors in both approaches further suggest that local demand effects play an important role in these isolated areas: there is a much larger decrease in share of workers in informal non-agriculture, whose products are mostly non-tradable. Consistent with this hypothesis, Appendix Table D7 shows that hot temperatures significantly decrease household consumption and nonfood consumption in particular in those remote areas.

[49] For details in the construction of this measure and its caveat, see Appendix B3.

Table 1: Wet-bulb Temperature and Sectoral Labor Share by Trade Openness

Trade openness is proxied by distance to the nearest major seaport

	Panel A: Panel Approach			
	(1) Agriculture	(2) Formal Non-Agriculture	(3) Informal Non-Agriculture	(4) Inactive and Unemployed
DD27 × (Open=0) (N)	0.0027	0.0002	-0.0027	-0.0000
	(0.0021)	(0.0009)	(0.0016)	(0.0005)
.	[0.0037]	[0.0012]	[0.0018]	[0.0018]
DD27 × (Open=1) (T)	-0.0486	0.0224	0.0307	-0.0047
	(0.0133)	(0.0081)	(0.0116)	(0.0066)
.	[0.0157]	[0.0069]	[0.0119]	[0.0055]
p-value (N) = (T)	0.0004	0.0086	0.0060	0.4795
Observations	1707	1707	1707	1707
Province × Age Group FE	x	x	x	x
Region × Age Group × Year FE	x	x	x	x
	Panel B: Long Differences Approach			
	(1) Agriculture	(2) Formal Non-Agriculture	(3) Informal Non-Agriculture	(4) Inactive and Unemployed
DD27 × (Open=0) (N)	0.0309	-0.0133	-0.0217	0.0041
	(0.0095)	(0.0077)	(0.0061)	(0.0035)
.	[0.0089]	[0.0067]	[0.0061]	[0.0045]
DD27 × (Open=1) (T)	-0.0812	0.0297	0.0648	-0.0134
	(0.0395)	(0.0198)	(0.0279)	(0.0166)
.	[0.0297]	[0.0115]	[0.0255]	[0.0099]
p-value (N) = (T)	0.0049	0.0381	0.0034	0.2942
Observations	156	156	156	156
Region × Age Group FE	x	x	x	x

Notes: Each panel presents the effects (as point change) of 1 SD increase in degree days above 27°C, separately for tradable and non-tradable markets, on sectoral employment shares. Unit of analysis is province-agegroup-year for the panel approach, and province-agegroup for the long differences approach. "Open" is an indicator that takes value 1 if the distance from a province centroid to the nearest major port is below the 70th percentile (approximately 200 km) and 0 otherwise. All regressions control for other weather variables (cold temperatures, second-order polynomials of precipitation and wind speed) and their interactions with the 'Open' dummy. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 150 km and and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Hot Temperatures Have Disproportionately Negative Effects on Agricultural Labor Supply. I test whether hot temperatures affect labor supply by estimating model (11) where the dependent variable is the number of hours worked in each 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 2 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 hours worked, conditional on working in a sector. In Panel B, the dependent variable is the mean hours worked of individuals in an analysis unit, regardless of whether an individual worked in a specific sector or not (i.e., individuals not working in a specific sector are considered as working zero hours).

Table 2: Wet-bulb Temperature and Labor Supply

Hours worked are computed for the primary and secondary jobs

	Panel A: Conditional Hours of Work (Intensive Margin)			
	(1)	(2)	(3)	
	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture	
DD27	-30.26	2.63	-4.76	
	(9.19)	(5.67)	(4.28)	
.	[8.97]	[4.43]	[3.03]	
Mean Outcome	1278	1800	1788	
	Panel B: Unconditional Hours of Work (Extensive Margin)			
	(1)	(2)	(3)	(4)
	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture	Total
DD27	-19.64	6.13	4.77	-8.87
	(4.46)	(2.63)	(2.41)	(5.98)
.	[5.15]	[2.20]	[2.44]	[4.71]
Mean Outcome	766	420	670	1864
Observations	1551	1551	1551	1551
Province × Age Group FE	x	x	x	x
Region × Age Group × Year FE	x	x	x	x

Notes: Each panel presents the effects (as number of hours change) of 1 unit increase in degree days above 27°C on hours of work last 12 months. 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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. All regressions use sampling weights. SOURCES: Data from Household Living Standards Survey 1992-1998, 2004-2018.

The results in Table 2 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. 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 Disproportionately Negatively Affect 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 productivity.[50] Table 3 shows that cold temperatures virtually do not affect sectoral labor productivity, while 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% decrease in revenue per worker

[50] 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 using a panel approach regression of the following form:

$$\ln w_{prt} = f(\mathrm{WBT}_{pt}) + g(\mathrm{R}_{pt}) + \gamma_p + \gamma_{rt} + \epsilon_{prt} \tag{15}$$

where $\ln w_{prt}$ 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. Revenue in agriculture is not restricted to crop production but also includes revenues from livestock, aquaculture, farm service, and forestry. 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 150 km and temporal correlation up to five lags. 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 E. Data limitations, however, do not allow me to estimate the marginal product of labor. Details on the construction of this analysis dataset can be found in Appendix B1.1.

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 D8 further shows that hot temperatures also negatively affect agricultural yields but the effects are smaller in magnitude relative to revenue per worker: one degree day higher than 27°C leads to approximately 0.4% decrease in yields of rice–the main staple crop. Given that the planting area is largely not affected by hot temperatures (Appendix Table D9), these findings suggest that heat's impact on agricultural labor transcend the commonly studied land productivity mechanism wherein lower crop yields drive labor reallocation out of agriculture.[51]

Table 3: Wet-bulb Temperature and Sectoral Labor Productivity: 2002-2016

	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture
	(1)	(2)	(3)
DD9	0.0000	-0.0047	0.0007
	(0.0006)	(0.0031)	(0.0025)
.	[0.0005]	[0.0033]	[0.0024]
DD27	-0.0105	0.0078	-0.0018
	(0.0039)	(0.0109)	(0.0071)
.	[0.0030]	[0.0088]	[0.0074]
Mean Outcome	2.28	5.33	3.49
Province FE	x	x	x
Region-by-Year FE	x	x	x
Observations	416	416	416

Notes: This table presents the effect of 1 unit increase in degree days above 27°C. Unit of analysis is province-year. Dependent variables are log of annual revenue per worker (2010 million VND). Mean outcomes are in log points. Agricultural revenues cover crop production, livestock, aquaculture, and forestry. All columns control for second-order polynomials of precipitation, and 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 150 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. SOURCES: Data from Vietnam Enterprise Census and Household Living Standards Survey 2002-2016.

Although hot temperatures do not affect non-agricultural labor productivity on average, a subset of non-agricultural workers is also adversely affected. Appendix Table D10 presents additional results from a model of firm-level fixed effects, using an unbalanced 15-year longitudinal dataset of firms that appeared at least twice during the

[51] That temperatures do not significantly affect cultivated land is not surprising given the agricultural land protection policies in Vietnam. In an effort to ensure food security, the government has issued multiple documents (e.g., Decree No.63/NQ-CP, Decision 124/QD-TTg, Resolution 17/2011/QH13) emphasizing agricultural/rice land protection with very few exemptions for land conversion to other non-agricultural purposes that serve the national and public interests.

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. The comparable effect for old construction firms is a reduction of about 0.6%. These findings support 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 in most Vietnamese provinces. 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 in Section 4.2 implies that the labor productivity mechanism likely holds in the long term as well. In less connected areas, however, the temperature effect on sectoral labor allocation runs in the opposite direction, suggesting the dominance of local demand effects as in the case of Indian districts documented by Liu, Shamdasani, and Taraz (2023). These findings are consistent with the two predictions of the theoretical model: hot temperature-induced negative shock on agricultural productivity increases the non-agricultural employment share in the case of small open economies where prices are exogenous, but reduces the non-agricultural labor share where prices are endogenously determined by production and demand forces in closed economies.

5.2 Differential Effects by Age Group and Sector of In-Migration

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. The theoretical model predicts that the reallocation of labor from agriculture to non-agriculture is decreasing in cost of working in the non-agricultural sector in the case of small open economies. Furthermore, this cost can be inferred from the observed gains in earnings among the sample of workers who switched sectors.

In this section, I provide evidence that the cost of switching from agriculture to informal non-agriculture is similar across age groups, but it is significantly more costly for older individuals to work a formal non-agriculture job. To do so, I estimate the following regression, using the sample of sector-switchers from a pool of three-consecutive-survey-wave individual panel data sets over the period 2002-2018:

$$\ln w_{jt} = \sum_a \sum_s \kappa_{a,s} \mathbb{I}_a \mathbb{I}_s + \psi U_{jt} + \gamma_j + \gamma_t + \epsilon_{jt} \tag{16}$$

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, and household farm or non-farm net profits.[52] 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 U.

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 $\kappa_{a,s\in\{f,i\}}$ and $\kappa_{a,g}$ 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 (16), 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.[53]

Results from this exercise, reported in Table 4, suggest two key findings. First, there are large gains in annual average earnings for workers who switched from agriculture to non-agricultural sectors, even after controlling for hours worked and individual-specific time-invariant unobserved characteristics (Column 1). 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 (Columns 2-3). These estimates are in line with evidence in Hamory et al. (2021), who suggest

[52] 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.

[53] 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.

an approximate gain of about 22% for individuals moving from agriculture to non-agriculture in Indonesia. When 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 (Columns 5-6).[54]

Second, across the two specifications, there is no differential gains across age groups when moving from agriculture to informal non-agriculture (Columns 3 and 6). However, the group of older workers (age 55-64) experience largest gains when switching into formal non-agriculture, whereas there is little evidence of differential gains between the younger two groups (Columns 2 and 5).

Viewed through the lens of the model, these findings imply that for workers starting in agriculture, the cost of switching into informal non-agriculture is lower than formal one, which explains the larger effects of hot temperatures in magnitude on the share of labor in informal non-agriculture compared to the formal non-agricultural sector (Table 1). Furthermore, the older the workers, the larger the switching costs they incur if getting a job in formal non-agriculture, which explains why the younger workers comprise most of those who shift into this sector in response to extreme temperatures.

5.3 Reconciling the Impacts of Temperature on Sectoral Labor Reallocation in the Short and Long terms

The analysis on potential mechanisms so far speaks to the effects of hot temperatures on sectoral employment shares, the relative temperature effects on formal and informal non-agriculture, as well as the heterogeneous temperature effects across age groups. I have not made any explicit argument on the relative larger temperature effects (in magnitude) on agricultural and non-agricultural employment shares when using the long differences approach compared to the panel one.

It is useful to recognize that the source of variation in the panel approach makes

[54] 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 D4). 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 4. 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.

Table 4: Gains in Labor Earnings among Switchers

Reference category is working in agriculture

	log earnings including profits constructed by hour worked					
	(1) Non Agriculture	(2) Formal	(3) Informal	(4) Non Agriculture	(5) Formal	(6) Informal
Age Group 24-39 (G1)	0.266	0.329	0.249	0.299	0.506	0.141
	(0.015)	(0.018)	(0.015)	(0.120)	(0.125)	(0.131)
Age Group 40-54 (G2)	0.236	0.304	0.222	0.279	0.505	0.125
	(0.015)	(0.019)	(0.015)	(0.121)	(0.125)	(0.131)
Age Group 55-64 (G3)	0.310	0.454	0.270	0.355	0.658	0.190
	(0.033)	(0.048)	(0.035)	(0.124)	(0.132)	(0.135)
p-value G1 = G2 = G3	0.0620	0.0123	0.2801	0.0812	0.0079	0.2068
p-value G1 = G2		0.3158	0.2029		0.9626	0.4403
p-value G1 = G3		0.0145	0.5666		0.0029	0.1966
p-value G2 = G3		0.0032	0.2021		0.0023	0.0810
Adj.R^2	0.544	0.545	0.545	0.548	0.550	0.550
Observations	37699	37699	37699	37699	37699	37699
Individuals	14053	14053	14053	14053	14053	14053
Year FE	x	x	x	x	x	x
Individual FE	x	x	x	x	x	x
Controls × Sector				x	x	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. Results for formal and informal non-agricultural sectors (Columns 2-3, Columns 5-6) are estimated jointly. Earnings include labor wages, other benefits and household farm/non-farm profits, trimmed at its top and bottom 5%. 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. 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 temperature-induced reallocation effects responses to an unanticipated shock. In this case, the higher costs associated with frictions are, the slower the flow of labor is. As a result, given that the cost of switching from agriculture to informal non-agriculture is lower than the switching cost to formal non-agriculture, the panel approach yields a larger temperature effect on informal non-agricultural labor share than on formal non-agricultural employment share.

When workers' career decision involves a dynamic discrete choice problem with recurring switching costs, however, what matters are the choices that forward-looking workers make in the face of the trend that global warming is disproportionately affecting agricultural productivity, making non-agricultural sectors relatively more attractive. The reason is that when workers face switching costs across sectors, career choices depend not only on current real wages, but also the career continuation values that reflect the option and associated costs of being employed in a particular sector. Under this scenario, hot temperatures will induce more forward-looking workers to move out of agriculture and into the non-agricultural sectors when evaluated over the longer time frame.

With this argument, if workers are indeed forward-looking and correctly assume that hotter places are expected to experience more damages from hot days relative to colder places in the face of global warming, then even in the short run, we should be able to see larger reallocation effects (in magnitude) out of agriculture in hotter relative to colder places. Indeed, results in Appendix Table D11 provide evidence that there are differential effects of short-run increase in hot temperatures on agricultural labor shares across hotter and less hot provinces. While extremely hot temperatures lower the share of the labor force engaging in agriculture across both hot and less hot areas, the point estimate for hot areas is significantly larger in magnitude and more precisely estimated than that for less hot areas. Similarly, extremely hot temperatures increase the share of workers in informal non-agriculture in both hot and less hot provinces, but the point estimate for hotter provinces is significantly larger and more precisely estimated than that for less hot provinces. For formal non-agriculture–the sector into which entry incurs higher cost, the effect is similar across the two groups of provinces.

In remote areas that are less integrated into global markets, on the other hand, the finding that extreme temperatures increase labor share in the agricultural sector more in the long run compare to the short run appears consistent with the intensification hypothesis in which liquidity constraints likely amplify the inability of workers and households to smooth consumption over time (Liu, Shamdasani, and Taraz 2023).

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 show that trade openness and labor market frictions play crucial roles in the temperature-sectoral labor reallocation relationship, and that extreme temperatures can affect sectoral employment choices beyond its well-documented negative impacts on crop productivity. These findings have important policy implications.

First, although climate change accelerating the reallocation of labor away from the relatively low-productivity agricultural sector in areas well-integrated to global markets may sound beneficial to the economy, the fact that it has negative impacts on labor productivity in other climate-exposed sectors, and that a nontrivial fraction of such reallocation is into the informal non-agricultural sector makes the overall effects of climate change-induced labor reallocation less certain. Due to labor market frictions that limit the movement of workers across sectors, informal non-agriculture does not appear an occupational pathway. There is little scope for agricultural workers to move to informal non-agriculture then subsequently move to formal non-agriculture.[55] Given that workers in the informal non-agricultural sector have lower education attainment and productivity, on average, if a large part of informal workers in the Vietnamese economy fall into the survival category–as in Brazil (Ulyssea 2018)–then climate change-induced labor 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).

Second, the opposite impacts of extreme temperatures on sectoral labor allocation between well-connected (where prices are less affected by temperatures) and isolated areas emphasize the role of trade openness to understanding the climate-employment relationship. This suggests that reducing trade barriers may be critical to mitigating climate damages in low and middle-income countries.

Finally, hot temperatures have differential effects on labor reallocation across age groups and sectors into which workers move. While workers of all age cohorts are equally likely to move into informal non-agriculture, younger workers comprise most of those who shift to a formal non-agricultural job. Given the nontrivial gaps in labor earnings between the informal and formal non-agricultural sectors, as well as the lack of access to social protection benefits for informal sector workers, these findings suggest that the older population likely disproportionately experiences work-related welfare impacts caused by climate change.

[55] Appendix Figure D5 shows that among individuals who worked in agriculture in the first period and in informal non-agriculture two years later, only 4.5% were able to take up a formal non-agricultural job in the third period. Once an agricultural worker was able to transition into the formal non-agricultural sector in the second period, however, they would face a 54% likelihood of continuing working in this sector two years later.

References

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.

Aragón, F. M., F. Oteiza, and J. P. Rud (2021). "Climate Change and Agriculture: Subsistence Farmers' Response to Extreme Heat". *American Economic Journal: Economic Policy* 13.1, pp. 1–35.

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 (2023). "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.

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.

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.

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.

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 (2012). “Temperature shocks and economic growth: Evidence from the last half century”. *American Economic Journal: Macroeconomics* 4.3, pp. 66–95.

— (2014). “What do we learn from the weather? The new climate-economy literature”. *Journal of Economic Literature* 52.3, pp. 2774–2814.

Donovan, K., W. J. Lu, and T. Schoellman (2023). “Labor Market Dynamics and Development”. *Quarterly Journal of Economics* qjad019.

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

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

Gollin, D., S. Parente, and R. Rogerson (2002). “The role of agriculture in development”. *American Economic Review* 92.2, pp. 160–164.

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. 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.

Hsiang, S. M. (2010). “Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America”. *Proceedings of the National Academy of Sciences* 107.35, pp. 15367–15372.

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*.

Jagnani, M., C. B. Barrett, Y. Liu, and L. You (2021). “Within-season producer response to warmer temperatures: Defensive investments by Kenyan farmers”. *Economic Journal* 131.633, pp. 392–419.

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.

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

La Porta, R. and A. Shleifer (2014). "Informality and development". *Journal of Economic Perspectives* 28.3, pp. 109–26.

Lagakos, D., S. Marshall, A. M. Mobarak, C. Vernot, and M. E. Waugh (2020). "Migration costs and observational returns to migration in the developing world". *Journal of Monetary Economics* 113, pp. 138–154.

Liu, M., Y. Shamdasani, and V. Taraz (2023). "Climate change and labor reallocation: Evidence from six decades of the Indian Census". *American Economic Journal: Economic Policy* 15.2, pp. 395–423.

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.

LoPalo, M. (2023). "Temperature, worker productivity, and adaptation: evidence from survey data production". *American Economic Journal: Applied Economics* 15.1, pp. 192–229.

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.

Matsuyama, K. (1992). "Agricultural productivity, comparative advantage, and economic growth". *Journal of Economic Theory* 58.2, pp. 317–334.

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.

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

Munshi, K. and M. Rosenzweig (2016). "Networks and misallocation: Insurance, migration, and the rural-urban wage gap". *American Economic Review* 106.01, pp. 46–98.

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.

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*. SSRN Working Paper http://dx.doi.org/10.2139/ssrn.4221478.

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

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.

Ulyssea, G. (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 a Developing Country

Trinh Pham

A Proofs of Theoretical Predictions

The share of employment and the efficiency unit of labor in agriculture and non-agriculture are determined as

$$
\begin{aligned}
L_g &= N_g = \int_1^{\hat{\varepsilon}} dF(\varepsilon) = 1 - \hat{\varepsilon}^{-\theta} = 1 - \left[\frac{pZ_gG'(L_g)}{(1-\tau)Z_nH'(L_n)}\right]^{-\theta} \\
L_n &= \int_{\hat{\varepsilon}} \varepsilon dF(\varepsilon) = \frac{\theta}{\theta-1}\hat{\varepsilon}^{1-\theta} \\
N_n &= \hat{\varepsilon}^{-\theta} = \left[\frac{pZ_gG'(L_g)}{(1-\tau)Z_nH'(L_n)}\right]^{-\theta}
\end{aligned}
$$

In the context of this study, hot temperatures affect sectoral labor reallocation through effects on sector-specific productivity Z_g and Z_n (and thus the supply of and demand for agricultural goods). While the adverse effect of heat on agricultural productivity is well-documented in the empirical literature (Schlenker and Roberts 2009), there is a dearth of evidence from developing settings on the relationship between heat and non-agricultural productivity, although existing evidence does suggest a negative heat impact on productivity of climate-exposed industries (Somanathan et al. 2021; LoPalo 2023).

Prediction 1: Small Open Economy. If the economy is sufficiently close to a small open economy in the absence of trade barriers and extreme temperatures disproportionately affect agricultural productivity, then

(a) extreme temperatures reduce the employment share of agriculture

(b) extreme temperatures increase the employment share of non-agriculture

(c) the reallocation effect induced by extreme temperatures is decreasing in the cost of working in non-agriculture

Proof. When the economy is sufficiently close to a small open economy, the relative agricultural price is held fixed by the world market $p = \bar{p}$. Temperature change affects sectoral labor supply though its effects on productivity.

$$N_g = L_g = 1 - \left[\frac{\bar{p}}{(1-\tau)Z_n}\right]^{-\theta}\left[\frac{Z_gG'(L_g)}{H'(L_n)}\right]^{-\theta} \tag{S1}$$

For simplicity, I assume (and later verify in the main analysis) that the negative temperature effect on non-agriculture is small and close to zero, and thus what matters is the effect of temperatures on agricultural productivity. Take implicit derivative of equation (S1) with respect to Z_g and rearranging terms

$$\frac{\partial N_g}{\partial Z_g} = \frac{\theta}{Z_g}\left[\frac{\Gamma}{(1-\tau)}\right]^{-\theta}\left\{1-\theta\left[\frac{\Gamma}{(1-\tau)}\right]^{-\theta}\left[\frac{G''(L_g)}{G'(L_g)}+\frac{\Gamma H''(L_n)}{(1-\tau)H'(L_n)}\right]\right\}^{-1} \tag{S2}$$
$$\text{where} \quad \Gamma := \frac{\bar{p}Z_g G'(L_g)}{Z_n H'(L_n)}$$

Because $G'(.) > 0, H'(.) > 0, G''(.) < 0, H''(.) < 0, 0 < \tau < 1, \theta > 0$, the right-hand side of equation (S2) is strictly positive. Therefore, $\frac{\partial N_g}{\partial Z_g} > 0$, which implies that if extreme temperatures cause a negative shock to agricultural productivity, then they decrease the employment share in agriculture.

By the same argument, extreme temperatures-induced negative shocks to agricultural productivity increase the employment share in non-agriculture.

$$\frac{\partial N_n}{\partial Z_g} = -\frac{\theta}{Z_g}\left[\frac{\Gamma}{(1-\tau)}\right]^{-\theta}\left\{1-\theta\left[\frac{\Gamma}{(1-\tau)}\right]^{-\theta}\left[\frac{G''(L_g)}{G'(L_g)}+\frac{\Gamma H''(L_n)}{(1-\tau)H'(L_n)}\right]\right\}^{-1} \tag{S3}$$

In addition, as the cost of working in non-agriculture increases within the range $0 < \tau < 1$, the right-hand side of equation (S3) decreases, implying that the marginal effect of extreme temperatures on non-agricultural employment share is a decreasing function in the cost of working in this sector. This suggests that holding all other things constant, if workers of all age cohorts share the same level of cost in a non-agricultural sector, the marginal effects of temperatures on employment shares in this sector are similar across age cohorts. On the other hand, if older workers incur a larger cost, the marginal effect of temperatures on non-agricultural employment share of older workers is smaller than that of younger workers.

Prediction 2: Closed Economy. If the economy is sufficiently closed and extreme temperatures disproportionately affect agricultural productivity, then under certain conditions

(a) extreme temperatures increase the employment share of agriculture

(b) extreme temperatures decrease the employment share of non-agriculture

(c) the reallocation effect (in magnitude) induced by extreme temperatures is decreasing in the cost of working in non-agriculture

Proof. When the economy is a closed system, the relative agricultural price p is endogenously determined. Good markets equilibrium requires that $C_g = Q_g$ and $C_n = Q_n$. These combined with equations (1) and (4) imply that

$$\begin{aligned} Z_g G(L_g) &= \zeta + \frac{\alpha}{(1-\alpha)p} Z_n H(L_n) \\ \text{where} \quad p &= \frac{(1-\tau) Z_n H'(L_n)}{Z_g G'(L_g)} (1-L_g)^{-1/\theta} \end{aligned} \tag{S4}$$

Take implicit derivative of equation (S4) with respect to agricultural productivity

$$\begin{aligned} \frac{\partial N_g}{\partial Z_g} &= \frac{-\zeta}{Z_g^2} \left[G'(L_g) - \frac{\alpha}{(1-\alpha)(1-\tau)} \Phi \right]^{-1} \\ \frac{\partial N_n}{\partial Z_g} &= \frac{\zeta}{Z_g^2} \left[G'(L_g) - \frac{\alpha}{(1-\alpha)(1-\tau)} \Phi \right]^{-1} \end{aligned} \tag{S5}$$

$$\begin{aligned} \text{where} \quad \Phi = &\frac{G''(L_g) H(L_n)}{H'(L_n)} (1-L_g)^{1/\theta} - G'(L_g) \\ &- \frac{G'(L_g) H(L_n)}{\theta H'(L_n)} (1-L_g)^{(1-\theta)/\theta} + \frac{G'(L_g) H(L_n) H''(L_n)}{[H'(L_n)]^2} \end{aligned}$$

Because $G'(.) > 0, H'(.) > 0, G''(.) < 0, H''(.) < 0, 0 < \tau < 1, \theta > 0$, and $\zeta > 0$, N_g is a decreasing function in agricultural productivity. Adverse shocks to agricultural productivity induced by extreme temperatures pull workers toward this sector. Engel's law plays a crucial role in this case. If $\zeta = 0$, then the share of workers in agriculture is independent of Z_g and thus agricultural productivity has no effect on labor allocation. If $\zeta < 0$, that is, if the agricultural good is a luxury good, then a positive shock to agricultural productivity is associated with an increase in agricultural labor share.

By the same argument, as the cost of working in non-agriculture increases within the range $0 < \tau < 1$, $\frac{\partial N_n}{\partial Z_g}$ increases, which implies that the magnitude of the marginal effect of extreme temperatures on non-agricultural employment share is decreasing in the cost of working in this sector.

B Data and Measurement

B1 Data

B1.1 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.

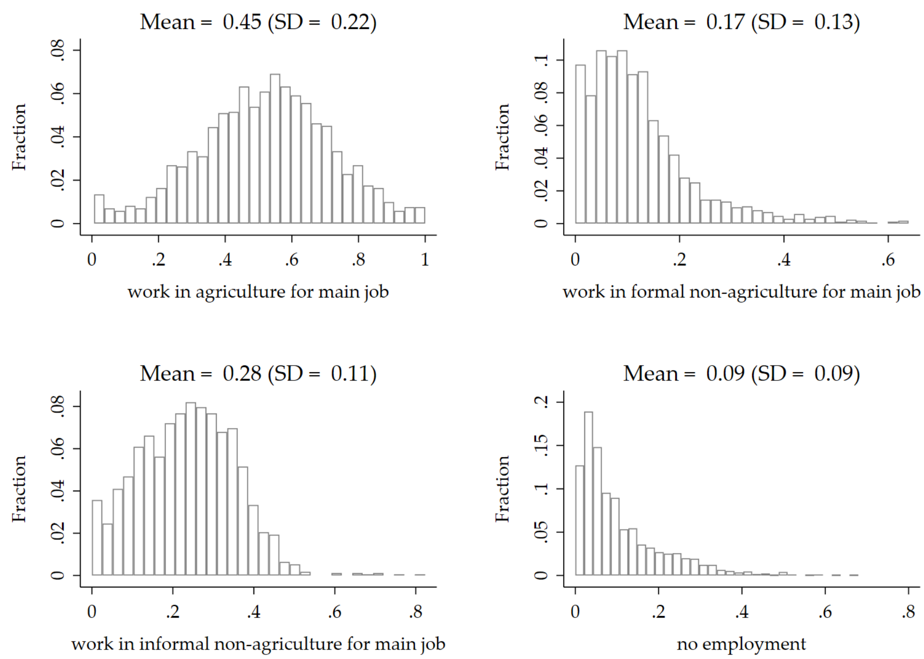
Table B1: List of Vietnamese Data Sets

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 Standards Survey 2002	12 months	30,000 households	GSO	Restricted
Household Living Standards Survey 2004	12 months	45,000 households	GSO	Restricted
Household Living Standards Survey 2006	12 months	45,000 households	GSO	Restricted
Household Living Standards Survey 2008	12 months	45,000 households	GSO	Restricted
Household Living Standards Survey 2010	12 months	45,000 households	GSO	Restricted
Household Living Standards Survey 2012	12 months	45,000 households	GSO	Restricted
Household Living Standards Survey 2014	12 months	70,000 households	GSO	Restricted
Household Living Standards Survey 2016	12 months	45,000 households	GSO	Restricted
Household Living Standards 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 B1: 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 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]

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

[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.

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 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 B2 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.

B1.2 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 2023).

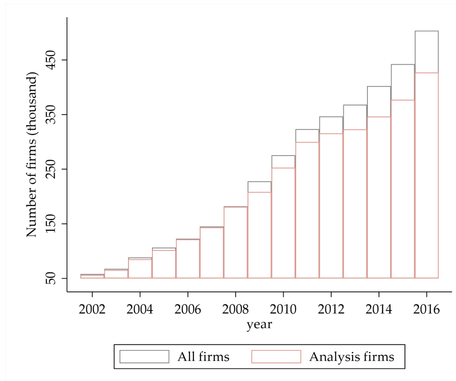
To calculate daily average WBT, I proceed in three steps. First, I calculate hourly relative humidity RH, which is defined as the ratio of vapor pressure e and saturation vapor pressure e_s, using hourly air temperature T_a (°C) and hourly dew-point temperature T_d (°C), following Bolton (1980):[3]

$$RH = 100 \times \frac{e}{e_s} = 100 \times \exp\left[\frac{17.67 \times 243.5 \times (\mathrm{T}_d - \mathrm{T}_a)}{(243.5 + \mathrm{T}_a)(243.5 + \mathrm{T}_d)}\right] \tag{S6}$$

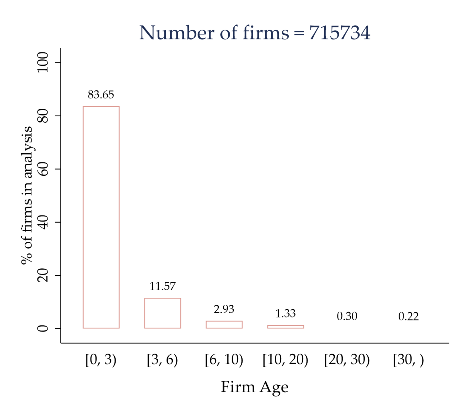
[3]This equation is available as archive on University Corporation for Atmospheric Research's website.

Figure B2: Non-agricultural firm-level data

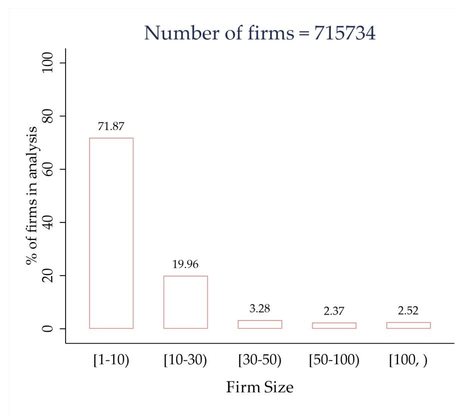
Panel A: Number of Firms



Panel B: Age Distribution



(a) Panel C: Size Distribution



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

$$\begin{aligned} \text{WBT} = \quad & T_a \times atan\left[0.151977 \times (RH + 8.313659)^{0.5}\right] + atan(T_a + RH) - 4.686035 \\ & - atan(RH - 1.676331) + 0.00391838(RH)^{1.5} \times atan(0.023101 \times RH) \end{aligned} \tag{S7}$$

Finally, I take the mean of hourly T_w to get daily WBT.

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.

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:

$$SPI_{pmy} = \frac{R_{pmy} - \bar{R}_{pm}}{\sigma_{pm}} \tag{S8}$$

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]

$$\text{wind speed} = \sqrt{U \times U + V \times V} \tag{S9}$$

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, nonlinear 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

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

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.

B1.3 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]

B2 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

[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.

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 B2: 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 is 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 B2 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 B3 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.

B3 Trade Openness Measurement

In the main analysis, I use the distance between a province's geographic centroid to the major ports as a proxy for market integration. In this section, I construct a different measure of trade openness, relying on price information in the household surveys. In particular, I proceed in four steps.

Table B3: 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

Notes: Informal workers are defined as those who do not have social security benefits nor labor contract. *Source:* Data from VHLSS 2014, 2016, 2018.

(i) 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

[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.

coefficients.[8] Because the world price series is only available from December 2003 with missing data for a few months between 2006-2008, one should interpret this correlation with caution.

(v) For Ho Chi Minh City, the largest economic hub of the country and home to the major Sai Gon seaport, there are insufficient data points on rice production. I assign the maximum value of pairwise correlation coefficients from other provinces for this city. The results are similar with the exclusion of this city.

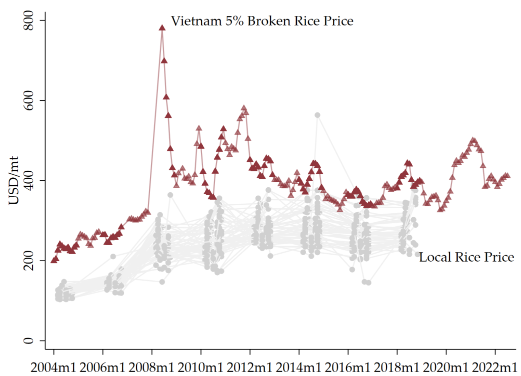
Figure B3 Panel A plot series of local and world rice prices over time. As seen, local rice prices are always lower than that of global markets and the two series share similar trends. Panel B shows that rice prices in most Vietnamese provinces are highly correlated with the world market price, with a mean value of correlation coefficient of 0.67. There is also significant variation across provinces. The maximum value of the correlation coefficient is nearly 0.9, whereas the minimum value is only 0.1. Panel C further shows that the two measures of trade openness (rice price coefficients and distance to the nearest major seaport), are strongly correlated.

I define a province as "open" if the distance from its geographic centroid to the nearest major seaport is below the 70th percentile (approximately 200 km). Table B4 presents the difference in returns to non-agricultural work, relative to agricultural work, across well-integrated and remote areas, which are obtained from estimating a variant of equation (16), using the sample of workers who changed sectors of employment. As seen, there is no statistical difference in the relative return to formal non-agricultural work across these two types of markets. If anything, the relative return to informal non-agriculture is significantly larger in remote areas less integrated to global markets than in other areas.

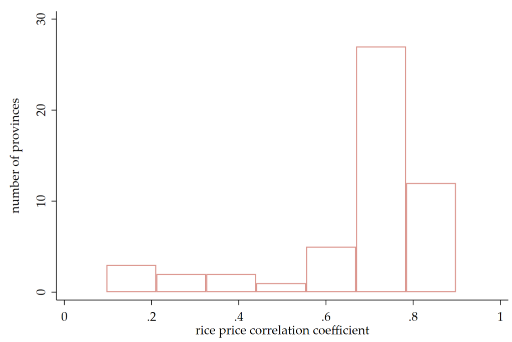
[8] The data on monthly world market price of Vietnamese rice 5% broken are obtained from the World Bank's "Pink Sheet," available at https://www.worldbank.org/en/research/commodity-markets.

Figure B3: Rice Price Series and Trade Openness Measures

Panel A: Local and World Rice Prices



Panel B: Distribution of provinces by rice price correlation coefficient



Panel C: Rice price correlations strongly correlated with distances to the neareast major seaport

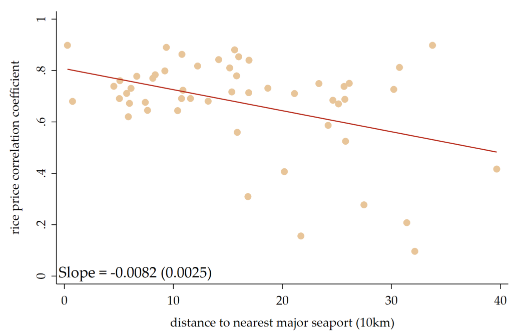


Table B4: Difference in Relative Returns to Non-Agricultural Employment between Well-integrated and Remote Areas

	Informal Non-Agriculture − Agriculture		Formal Non-Agriculture − Agriculture	
	(1)	(2)	(3)	(4)
Isolated Areas - Connected Areas	0.101	0.077	0.045	0.017
	(0.024)	(0.024)	(0.031)	(0.031)
Adj.R^2	0.545	0.550	0.545	0.550
Observations	37699	37699	37699	37699
Individuals	14053	14053	14053	14053
Year FE	x	x	x	x
Individual FE	x	x	x	x
Controls × Sector		x		x

Notes: This table presents difference in return to informal non-agricultural employment (columns 1-2) and return to formal non-agricultural employment (columns 3-4), relative to agricultural employment, between isolated and connected areas. Results for formal and informal non-agricultural sectors in each specification (columns 1 and 3, columns 2 and 4) are estimated jointly. 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%. 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. 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). Columns (1)-(3) are estimated jointly. 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.

C Robustness Checks and Additional Analyses

C1 Robustness

Robust to alternative specifications. Table C1 reports the results from estimating a variant of equation (11) in which the response functions $f(.)$ and $g(.)$ are not specific to any age group.

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 time trends, allowing different age groups to follow different trends in a limited way. In column (3), which is the preferred specification, age groups of different climatic and economic regions are assumed to follow more flexible trends. The estimates can be interpreted as the point change in sectoral employment share resulting from one SD increase in degree days below 9°C or higher than 27°C wet-bulb temperature during the 12-month reference period.

Robust to sample restrictions. In the baseline analysis, I keep all province-age group-year cells constructed from 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, and is arguably not representative at the province level. The effects are largely unchanged under the exclusion of those cells (Appendix Figures C2 and C3).

Similarly, to ensure that the results are not driven by some particular observations, I re-estimate the main specifications with all observations of a province being randomly dropped from each estimation. Figure C1 shows that the estimated coefficients remain stable and statistically significant under this test.

Robust to inclusion of time-varying demographic controls. 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. I test this concern by controlling for time-varying demographic characteristics that might influence sectoral employment, including educational attainment, share of male workers, and share of Kinh ethnic majority when estimating equations (11) and (13). 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

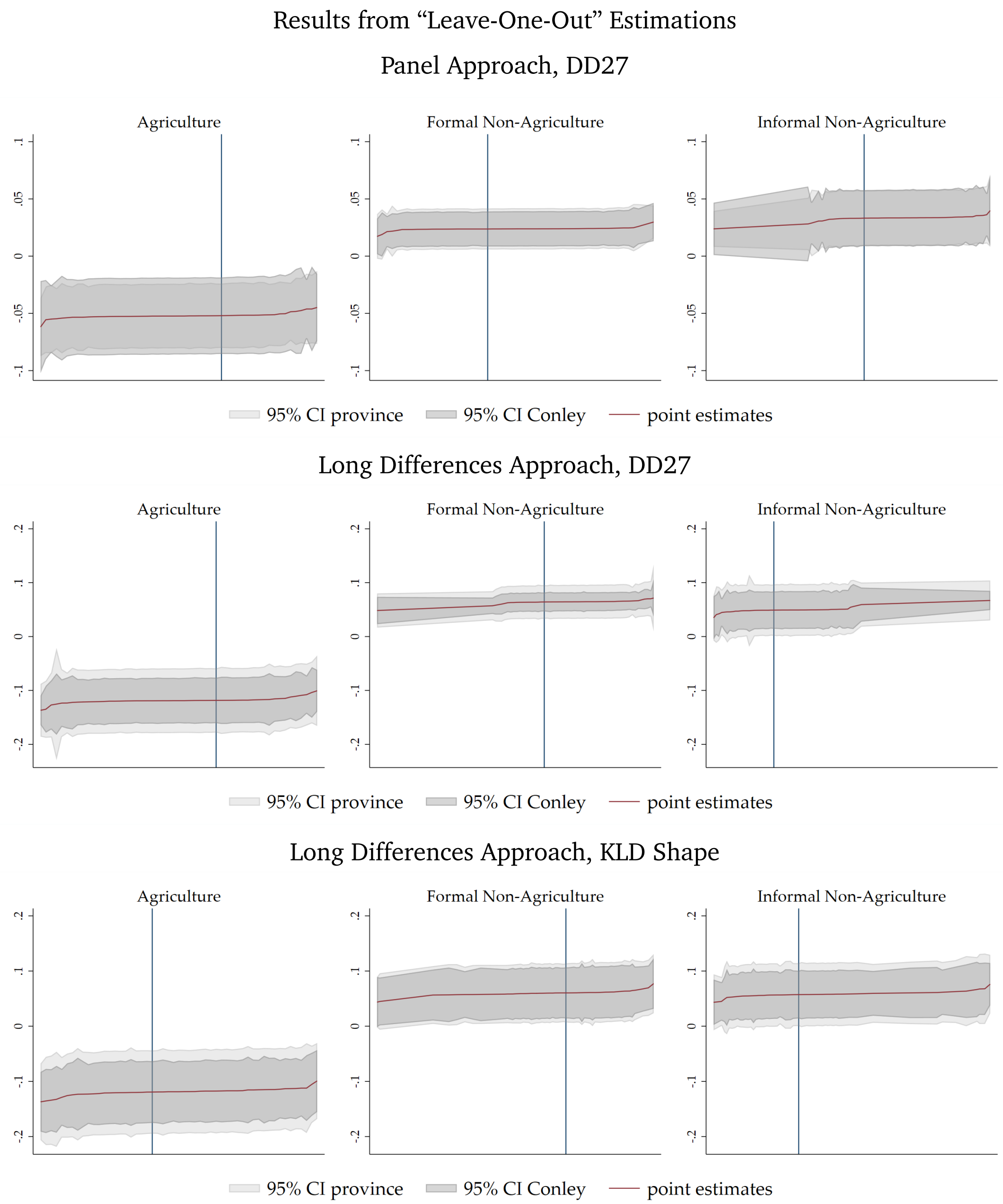
Table C1: Wet-bulb Temperature and Sectoral Labor Share, Panel Approach

Effects of 1 SD Increase in Weather Variables

	Agriculture			Formal Non-Agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)
DD9	-0.0016	-0.0019	-0.0026	-0.0010	-0.0007	-0.0011
	(0.0085)	(0.0086)	(0.0091)	(0.0036)	(0.0039)	(0.0040)
.	[0.0066]	[0.0067]	[0.0068]	[0.0034]	[0.0037]	[0.0036]
DD27	-0.0459	-0.0493	-0.0520	0.0178	0.0214	0.0237
	(0.0132)	(0.0133)	(0.0143)	(0.0101)	(0.0091)	(0.0091)
.	[0.0160]	[0.0166]	[0.0170]	[0.0073]	[0.0073]	[0.0078]
Precipitation	-0.0381	-0.0390	-0.0391	0.0111	0.0120	0.0120
	(0.0188)	(0.0188)	(0.0198)	(0.0093)	(0.0094)	(0.0098)
.	[0.0171]	[0.0170]	[0.0171]	[0.0084]	[0.0084]	[0.0084]
Precipitation squared	0.0163	0.0167	0.0162	-0.0016	-0.0016	-0.0013
	(0.0130)	(0.0129)	(0.0136)	(0.0054)	(0.0053)	(0.0055)
.	[0.0104]	[0.0103]	[0.0103]	[0.0050]	[0.0049]	[0.0048]
Mean Outcome	0.45	0.45	0.45	0.17	0.17	0.17
Adjusted R2	0.86	0.88	0.88	0.83	0.90	0.92
	Informal Non-Agriculture			Inactive and Unemployed		
	(1)	(2)	(3)	(4)	(5)	(6)
DD9	0.0018	0.0017	0.0025	-0.0003	-0.0002	-0.0001
	(0.0059)	(0.0059)	(0.0062)	(0.0009)	(0.0008)	(0.0009)
.	[0.0048]	[0.0048]	[0.0047]	[0.0010]	[0.0010]	[0.0011]
DD27	0.0332	0.0323	0.0332	-0.0052	-0.0046	-0.0051
	(0.0120)	(0.0120)	(0.0122)	(0.0069)	(0.0070)	(0.0071)
.	[0.0124]	[0.0124]	[0.0125]	[0.0057]	[0.0058]	[0.0058]
Precipitation	0.0232	0.0231	0.0231	0.0038	0.0039	0.0040
	(0.0159)	(0.0160)	(0.0166)	(0.0059)	(0.0059)	(0.0062)
.	[0.0131]	[0.0132]	[0.0131]	[0.0049]	[0.0050]	[0.0050]
Precipitation squared	-0.0126	-0.0128	-0.0128	-0.0020	-0.0021	-0.0020
	(0.0112)	(0.0112)	(0.0117)	(0.0031)	(0.0031)	(0.0032)
.	[0.0070]	[0.0070]	[0.0070]	[0.0027]	[0.0027]	[0.0028]
Mean Outcome	0.28	0.28	0.28	0.09	0.09	0.09
Adjusted R2	0.70	0.71	0.72	0.91	0.91	0.92
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 shares of employment in each sector. All columns also control for 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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. All regressions use sampling weights.

Figure C1: Sensitivity Test: Wet-bulb Temperature and Sectoral Labor Share



Notes: Each graph presents the distribution of estimated coefficients of temperature effects (DD27 or KLD Shape) on sectoral employment shares, which are obtained from reestimating the main specifications in which all observations of a province are randomly dropped from the sample. Vertical lines are baseline estimates.

considered outcome variables. In Appendix Figures C2, C3, and C4, I show that the coefficients on sectoral employment shares are of similar magnitude to those obtained from the baseline specifications for both panel and long differences approaches and are more precisely estimated. In some specifications, the new coefficients are of slightly smaller magnitude but remained highly statistically significant. These findings suggest that changes in these demographic characteristics cannot explain entirely for the changes in sectoral employment shares induced by temperatures.

Robust to controlling for lagged outcomes. Because sectoral employment shares at the local level are highly correlated from one year to the next, panel estimation from equation (11) 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 Appendix Figure C2, the coefficients obtained from such a dynamic panel model are generally of similar magnitude to the coefficients from the baseline specification.

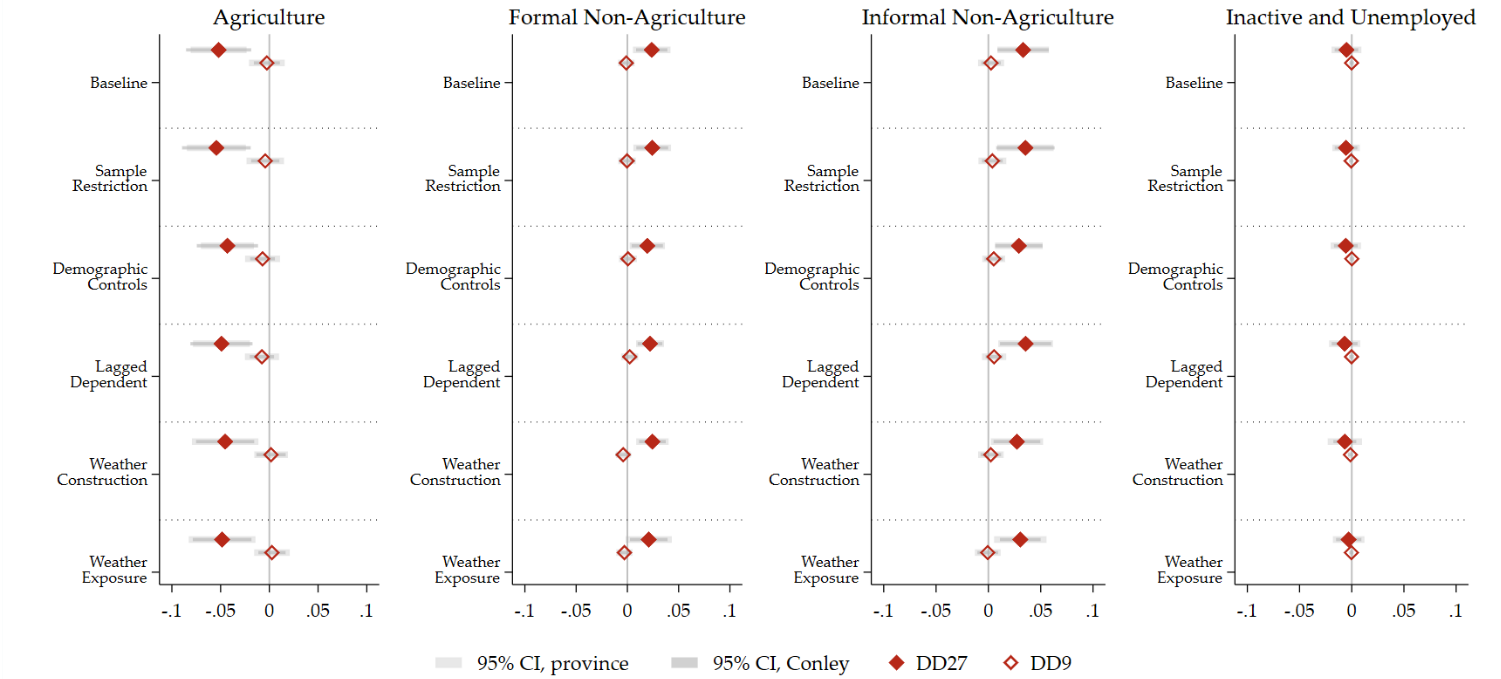
Robust to alternate method of constructing weather variables. In the baseline analysis, provincial level weather variables are computed as the weighted average of the four grid points closest to provincial geographic centroid, with weights being the inverse distance of weather grids to the province centroid. 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 (Appendix Figures C2, C3, and C4).

Robust to alternate definition of weather exposure. In merging the weather data with individual-level data, I assume that 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. The new estimates are of smaller magnitude than the corresponding baseline coefficients but remain statistically significant (Appendix Figures C2 and C3).

Robust to different period definitions (end points). 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: 1992-2008 and 2009-2018. The first period covering six household survey waves (1993, 1998, 2002, 2004, 2006, and 2008) and the second period covering five survey waves (2010, 2012, 2014, 2016, and 2018). In an alternate estimation, I divide the study period into two sub-periods: 1992-2007 (five survey waves) versus 2008-2018 (six survey waves), and two ten-year sub-periods (1992-2002 vs. 2008-2018). The effects on agricultural and formal non-agricultural employment shares are of similar magnitudes and remain statistical significance at the 5-10% level. The estimates for informal non-agriculture are positive but less precisely estimated. (Appendix Figure C3 and C4).

Figure C2: Robustness: Wet-bulb Temperature and Sectoral Labor Share

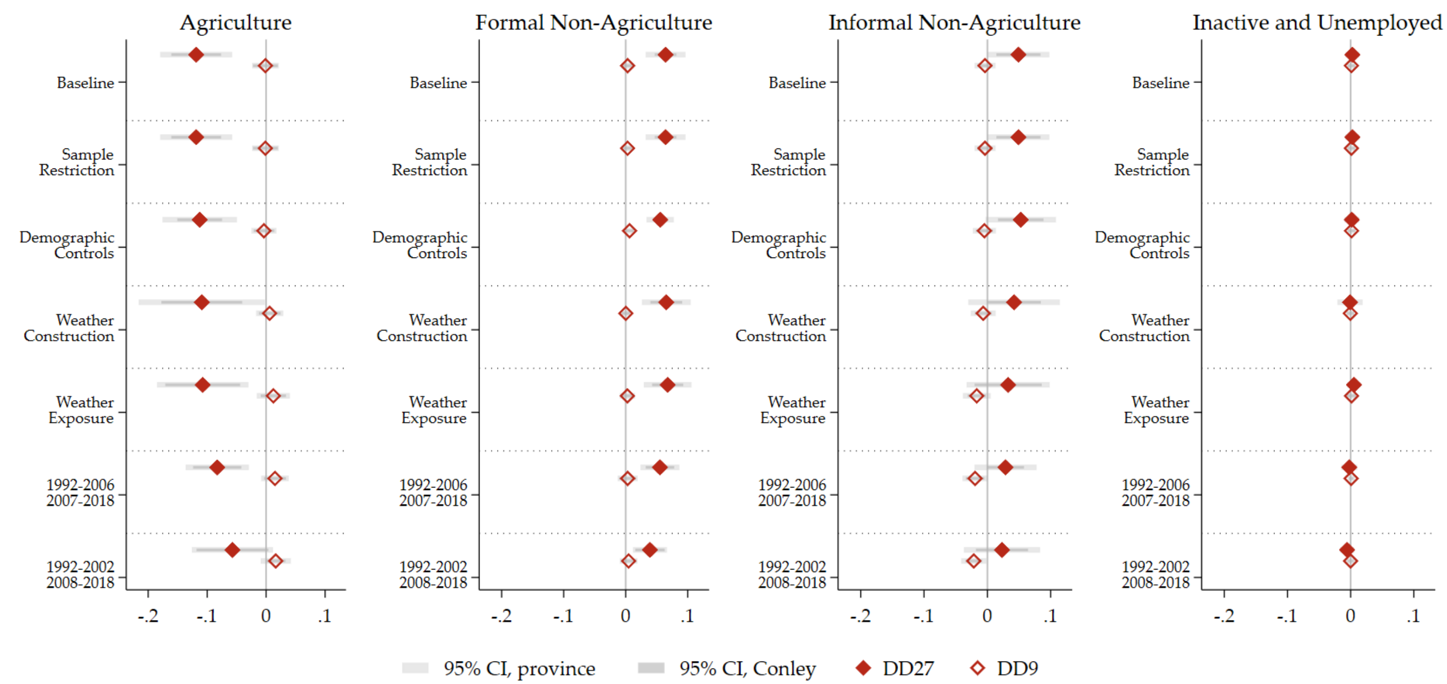
Results from Panel Approach with Degree Days



Notes: This figure presents the effects of 1 SD increase in hot wet-bulb temperatures above 27°C and cold temperatures below 9°C on sectoral employment shares, which are obtained from estimating equation (11). In the baseline specification, controls include province-by-age group and region-by-age group-by-year fixed effects, second-order polynomials of precipitation, number of days with high wind speeds during the 12-month exposure. In the "Sample Restriction" specification, province-age group-year cells constructed from less than 30 individual observations are excluded. In "Demographic Controls" specification, relative to the baseline, other time-varying demographic characteristics including educational attainment, share of male workers, and share of ethnic minority are also controlled. In the "Lagged DepVar" specification, relative to the baseline, the lagged dependent variables are controlled. In the "Weather Construction" specification, weather variables are constructed by taking average of all grid points within a geographic boundary, instead of weighted average of four nearest grid points to the geographic centroid. In the "Weather Exposure" specification, individuals are assumed to be exposed to 14-month weather distribution prior to the survey. Robust standard errors are clustered at the province level. Conley standard errors allow for spatial correlation up to 150 km and serial correlation up to five lags. Province distances are computed from province geographic centroids. All regressions use sampling weights. SOURCES: Data from Vietnam Household Living Standards Survey 1992-2018.

Figure C3: Robustness: Wet-bulb Temperature and Sectoral Labor Share

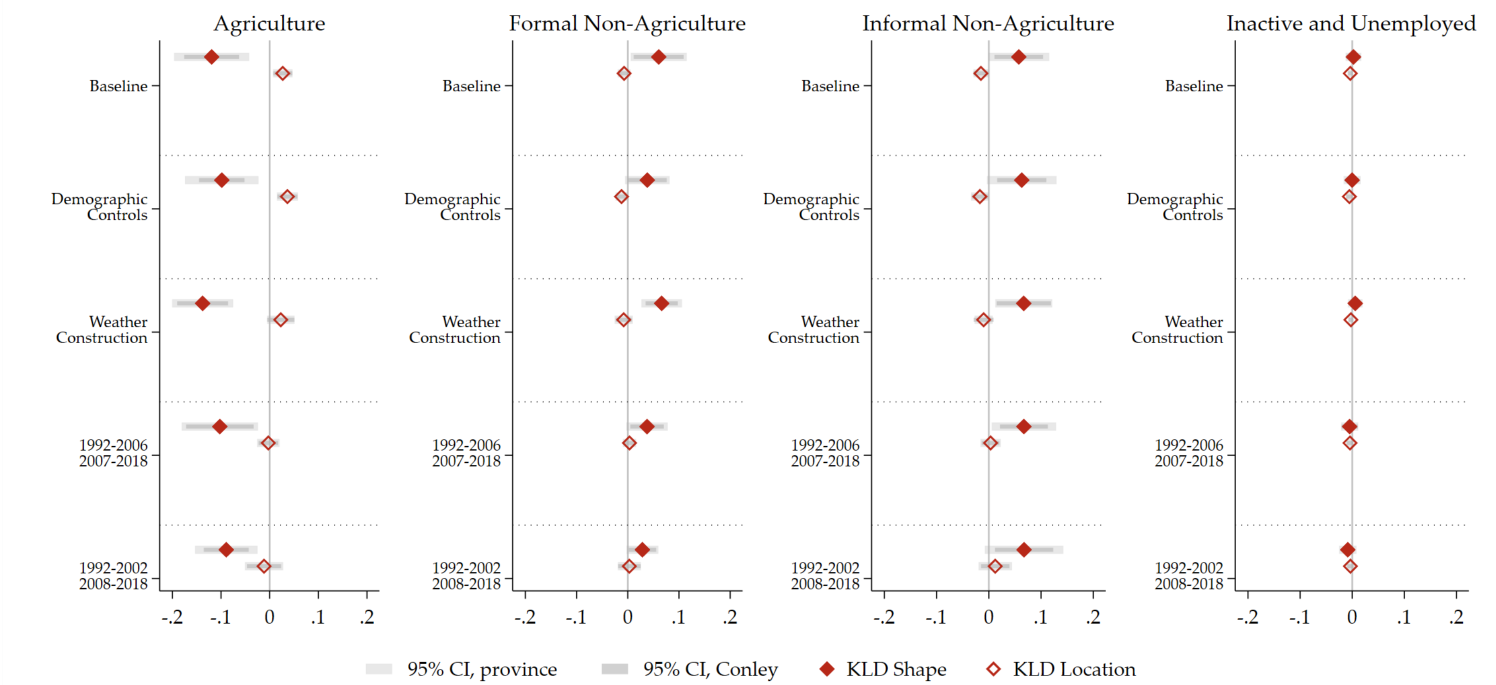
Results from Long Differences Approach with Degree Days



Notes: This figure presents the effects of hot temperatures on sectoral employment shares, which are obtained from estimating equation (13). In the baseline specification, I take the difference in outcomes (and weather variables) between two periods: 1992-2008 and 2009-2018, controls include region-by-age group fixed effects, 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 during the 12-month exposure. In the "Sample Restriction" specification, province-age group-year cells constructed from less than 30 individual observations are excluded. In the "Demographic Controls" specification, relative to the baseline, other time-varying demographic characteristics including educational attainment, share of male workers, and share of ethnic minority are also controlled. In the "Weather Construction" specification, weather variables are constructed by taking average of all grid points within a geographic boundary, instead of weighted average of four nearest grid points to the geographic centroid as in the baseline specification. In the "Weather Exposure" specification, individuals are assumed to be exposed to 14-month weather distribution prior to the survey. Robust standard errors are clustered at the province level. Conley standard errors allow for spatial correlation up to 150 km and serial correlation up to five lags. Province distances are computed from province geographic centroids. All regressions use average sampling weights.

Figure C4: Robustness: Wet-bulb Temperature and Sectoral Labor Share

Results from Long Differences Approach with KLD Measures



Notes: This figure presents the effects of changes in temperature distribution, as proxied by KLD shape and KLD location, on sectoral employment shares, which are obtained from estimating equation (13) and evaluated at the sample mean of the temperature variables. In the baseline specification, I take the difference in outcomes (and weather variables) between two periods: 1992-2008 and 2009-2018, controls include region-by-age group fixed effects, 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 during the 12-month exposure. In the "Demographic Controls" specification, relative to the baseline, other time-varying demographic characteristics including educational attainment, share of male workers, and share of ethnic minority are also controlled. In the "Weather Construction" specification, weather variables are constructed by taking average of all grid points within a geographic boundary, instead of weighted average of four nearest grid points to the geographic centroid as in the baseline specification. Robust standard errors are clustered at the province level. Conley standard errors allow for spatial correlation up to 150 km and serial correlation up to five lags. Province distances are computed from province geographic centroids. All regressions use average sampling weights.

Robust to alternative functional forms of temperatures. I estimate year-to-year panel models where the temperature function is represented by degree day bins, and fourth-order polynomials of daily average temperatures, summed across year. These models provide sufficient flexibility to capture important non-linearity, as well as being relatively parsimonious with low demand on the data.[9]

As for cumulative temperature bins, denote the endpoints of the eleven temperature bins (less than 9°C, 9 two-degree wide bins, higher than 27°C) by $[\mathrm{WBT}^{1k}, \mathrm{WBT}^{2k})$ with $k \in \{1, 2, ..., 11\}$, and assume that a day with WBT contributes positive degrees to the bin for which $\mathrm{WBT}^{1k} < \mathrm{WBT} \leq \mathrm{WBT}^{2k}$ and zero to all others. I assume that for $\mathrm{WBT} < 19$°C, the day contributes $\mathrm{WBT}^{2k} - \mathrm{WBT}$ to bin k; while for $\mathrm{WBT} > 19$°C, the day contributes $\mathrm{WBT} - \mathrm{WBT}^{1k}$ to bin k. These values are then summed across the reference period to determine the number of degrees in each bin.[10]

As for degree day bins, again denote the endpoints of the eleven temperature bins (less than 9°C, 9 two-degree wide bins, higher than 27°C) by $[\mathrm{WBT}^{1k}, \mathrm{WBT}^{2k})$ with $k \in \{1, 2, ..., 11\}$. I follow Somanathan et al. (2021) and consider a daily mean WBT contributes positive degrees to the bin for which $\mathrm{WBT} > \mathrm{WBT}^{1k}$ and zero to all others. If $\mathrm{WBT} \geq \mathrm{WBT}^{2k}$, the day contributes $\mathrm{WBT}^{2k} - \mathrm{WBT}^{1k}$ to bin k; if $\mathrm{WBT}^{1k} < \mathrm{WBT} \leq \mathrm{WBT}^{2k}$ then it contributes $\mathrm{WBT} - \mathrm{WBT}^{1k}$ to bin k.

As seen in Appendix Figures C5 and C6, the panel results from the baseline, parsimonious model are robust to alternative functional forms of temperatures: all the reallocation effects happen at the higher end of the temperature distribution–above approximately 27°.

[9]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 (11) 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 j in province p before collapsing into province-age group-year cell.

[10]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).

Figure C5: Wet-bulb Temperature and Sectoral Labor Share

Results from Panel Approach with Cumulative Temperature Bins

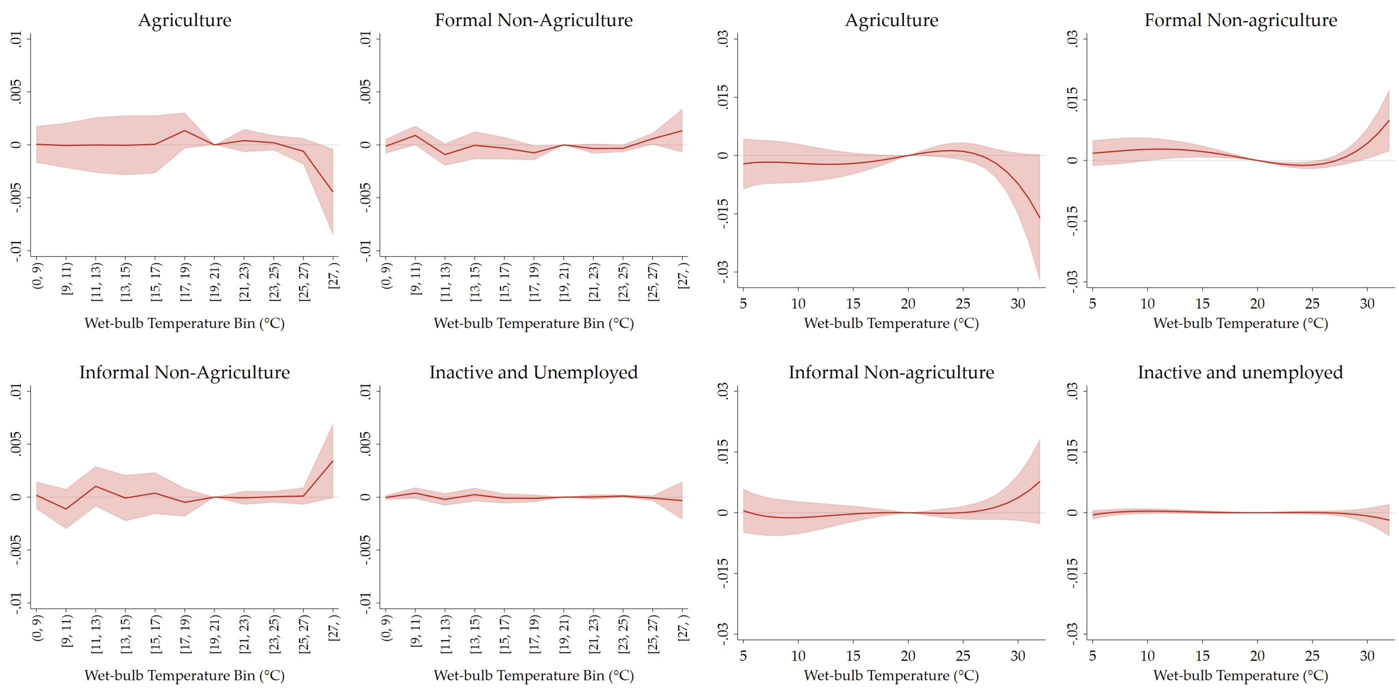
Notes: This figure shows that the relationship between wet-bulb temperature and primary sectoral employment share is robust to alternative functional form of temperatures. Each graph represents a predicted sectoral employment share-temperature response function (equation 11). Shaded areas are 95% confidence interval where robust standard errors are clustered at the province level.

Figure C6: Wet-bulb Temperature and Sectoral Labor Share

Results from Panel Approach

Panel A: Degree Day Bins

Panel B: Fourth-Order Polynomials

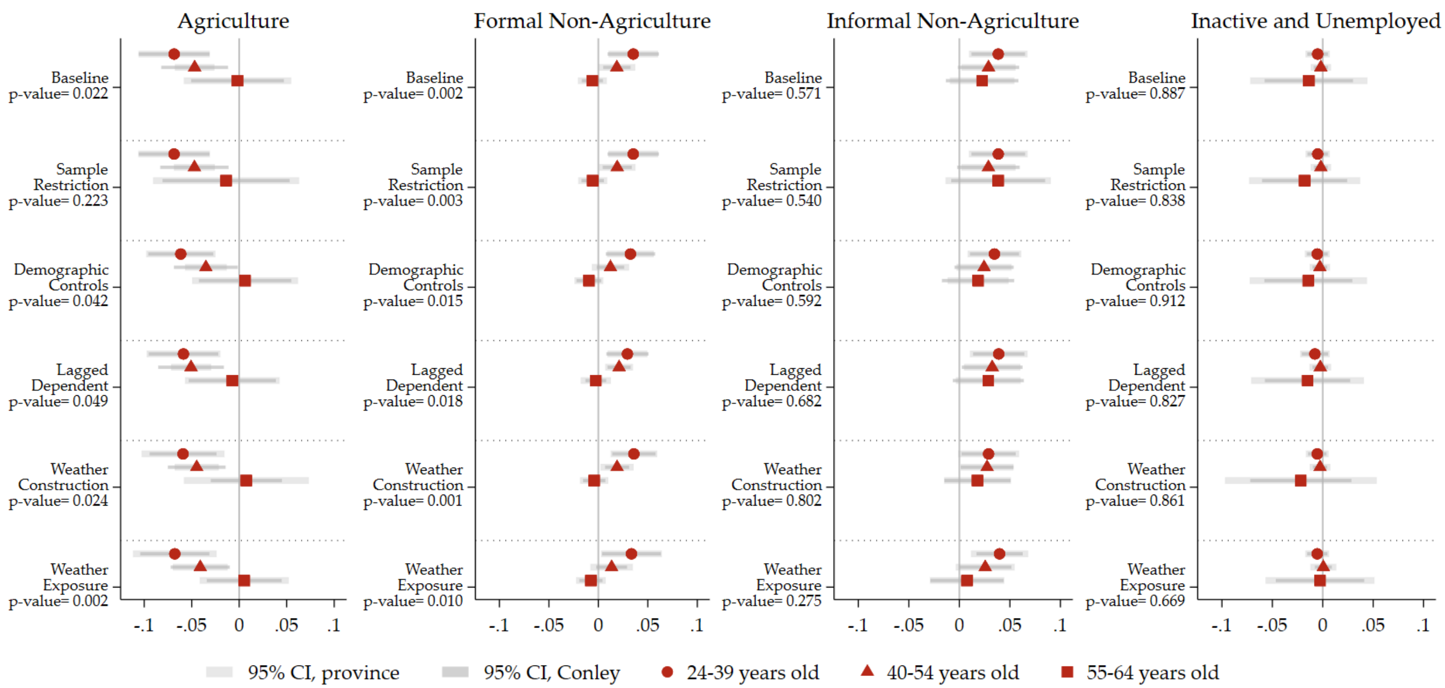


Notes: This figure shows that the relationship between wet-bulb temperature and primary sectoral employment share is robust to alternative functional forms of temperatures. Each graph represents a predicted sectoral employment share-temperature response function (equation 11). Shaded areas are 95% confidence interval where robust standard errors are clustered at the province level.

Heterogeneity results hold under these additional checks. Specifically, consistent with findings from the baseline specification, the effect of temperature change (hot days) on the formal non-agricultural labor share declines as one moves from the youngest to the oldest group. As for the informal non-agricultural employment share, however, the temperature effect is similar across the three groups. There is little evidence of temperature effect on non-employment across groups (Appendix Figures C7, C8, C9, C10, and C11).

Figure C7: Wet-bulb Temperature and Sectoral Labor Share by Age Group

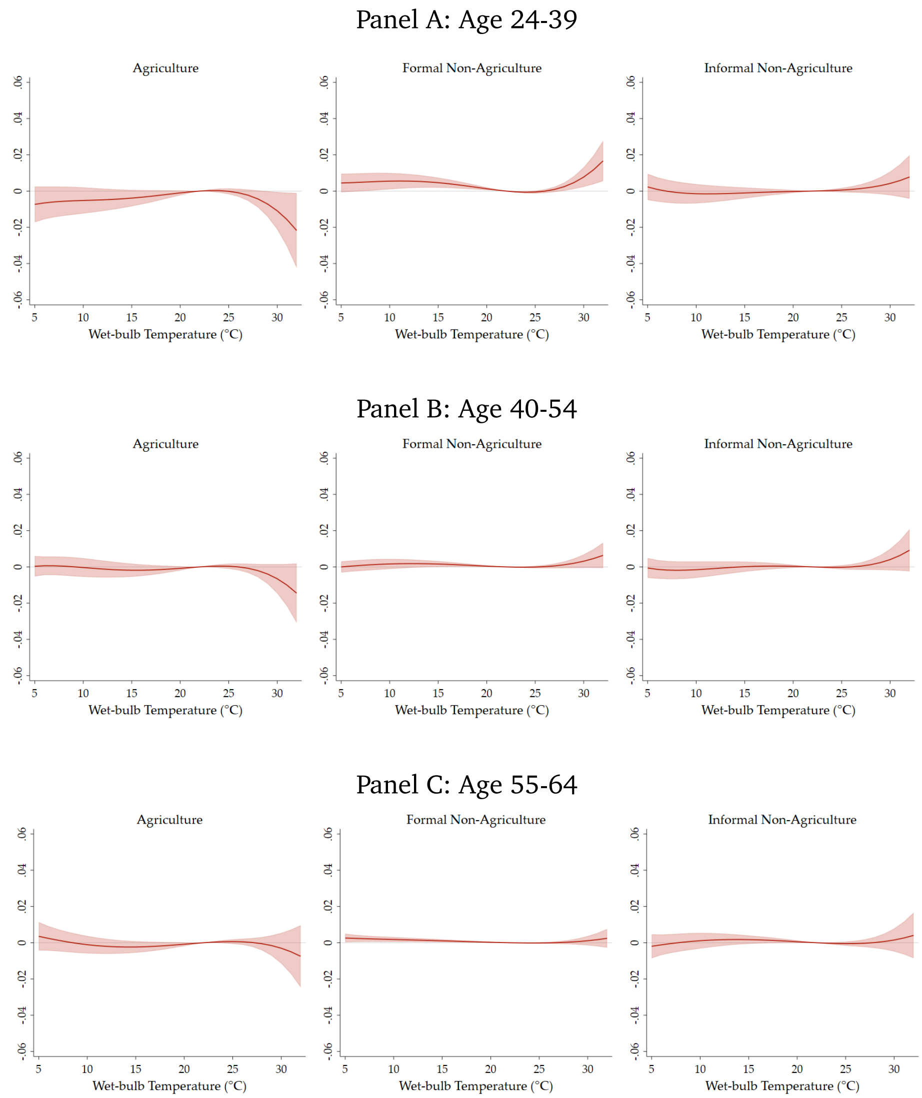
Results from Panel Approach with Degree Days Measures (DD27)



Notes: Results from estimating equation (11). Dependent variables are shares of employment in each sector. Province distances are computed from province geographic centroids. All regressions use sampling weights. p-values from the test of significant age cohort differences using standard errors clustered at the province level are reported. The results are qualitatively similar when using Conley standard errors.

Figure C8: Wet-bulb Temperature and Sectoral Labor Share by Age Group

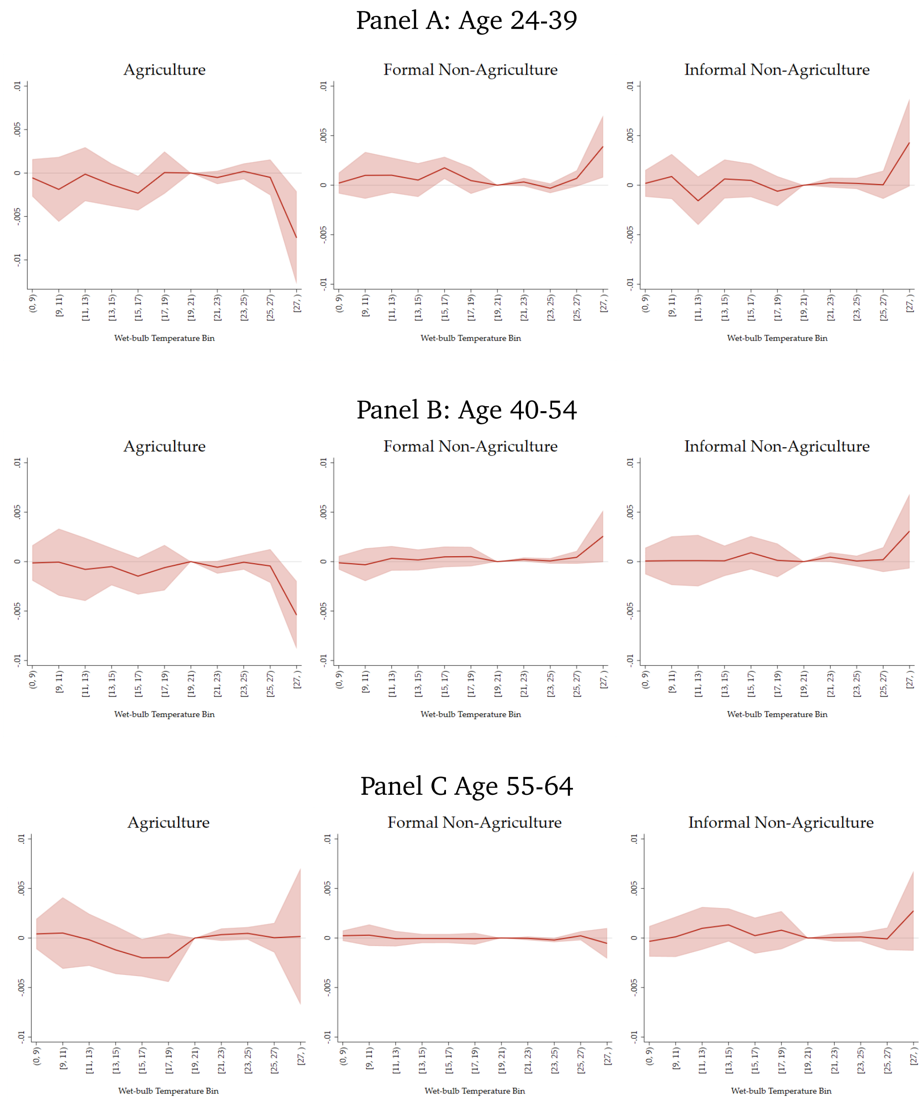
Results from Panel Approach with Fourth-Order Polynomials



Notes: Each graph represents a predicted sectoral employment share-temperature response function, estimated with equation (11). Shaded areas are 95% confidence interval where robust standard errors are clustered at the province level. Regression estimates are from a model of fourth-order polynomials in daily mean wet-bulb temperature. Details are in Section C1. 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.

Figure C9: Wet-bulb Temperature and Sectoral Labor Share by Age Group

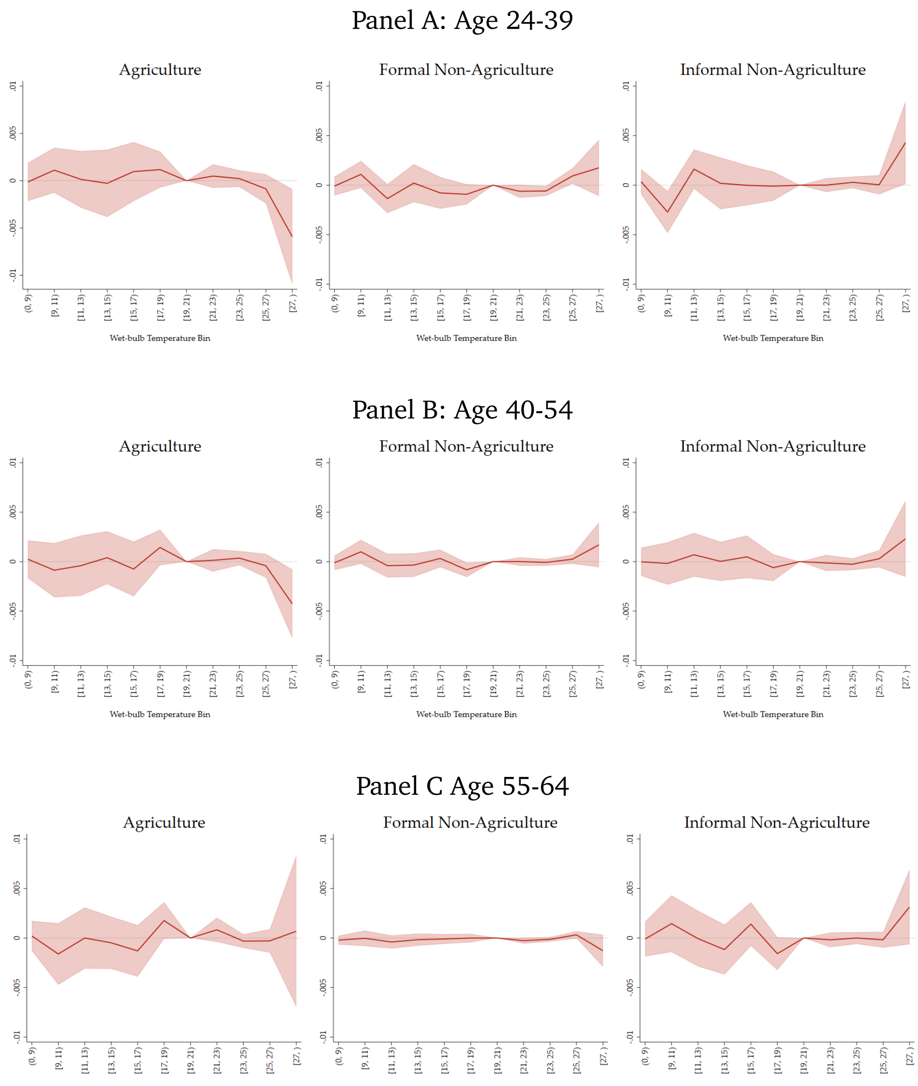
Results from Panel Approach with Cumulative Temperature Bins



Notes: Each graph represents a predicted sectoral employment share-temperature response function, estimated with equation (11). Shaded areas are 95% confidence interval where robust standard errors are clustered at the province level. Regression estimates are from a model of cumulative temperature bins. Details are in Section C1. 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.

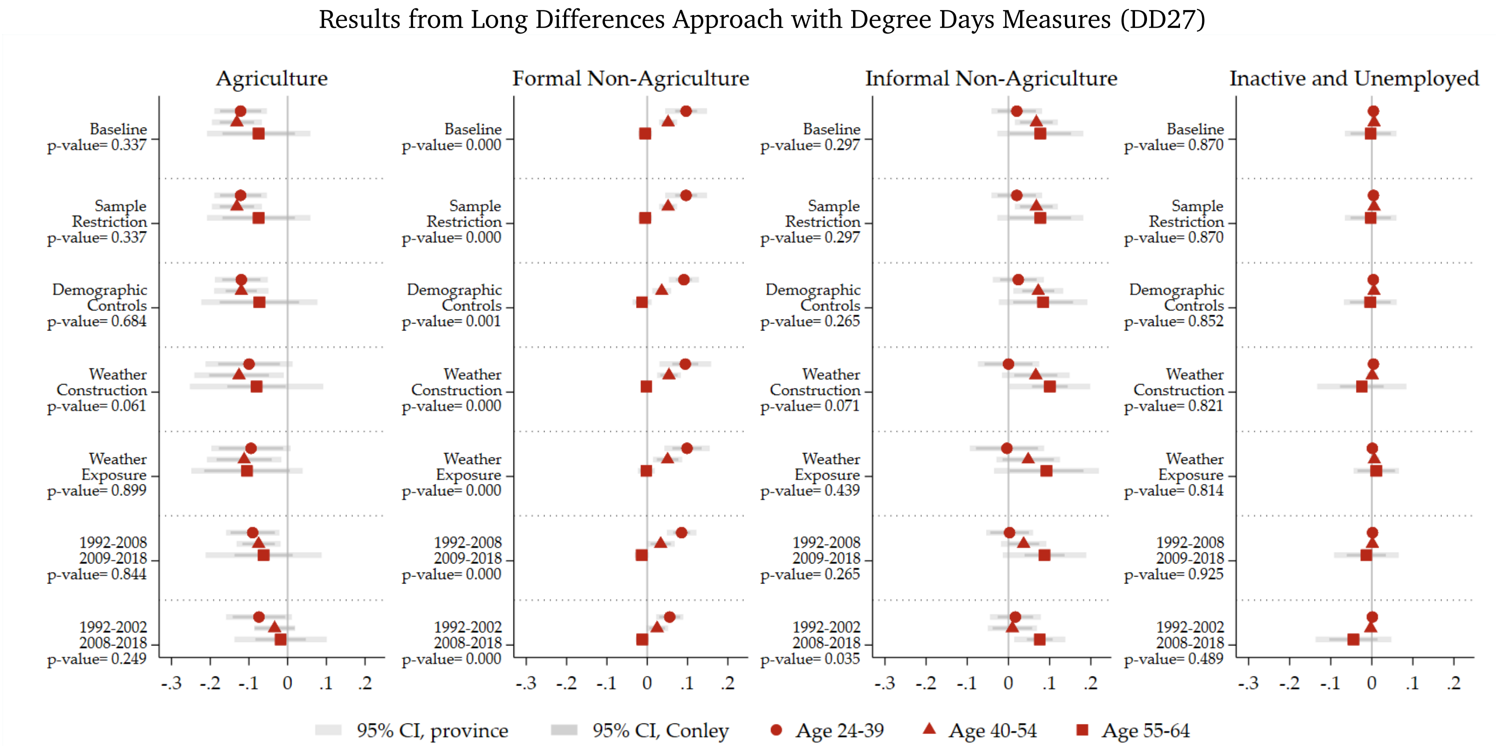
Figure C10: Wet-bulb Temperature and Sectoral Labor Share by Age Group

Results from Panel Approach with Degree Days Bins



Notes: Each graph represents a predicted sectoral employment share-temperature response function, estimated with equation (11). Shaded areas are 95% confidence interval where robust standard errors are clustered at the province level. Regression estimates are from a model of degree day bins. Details are in Section C1. 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.

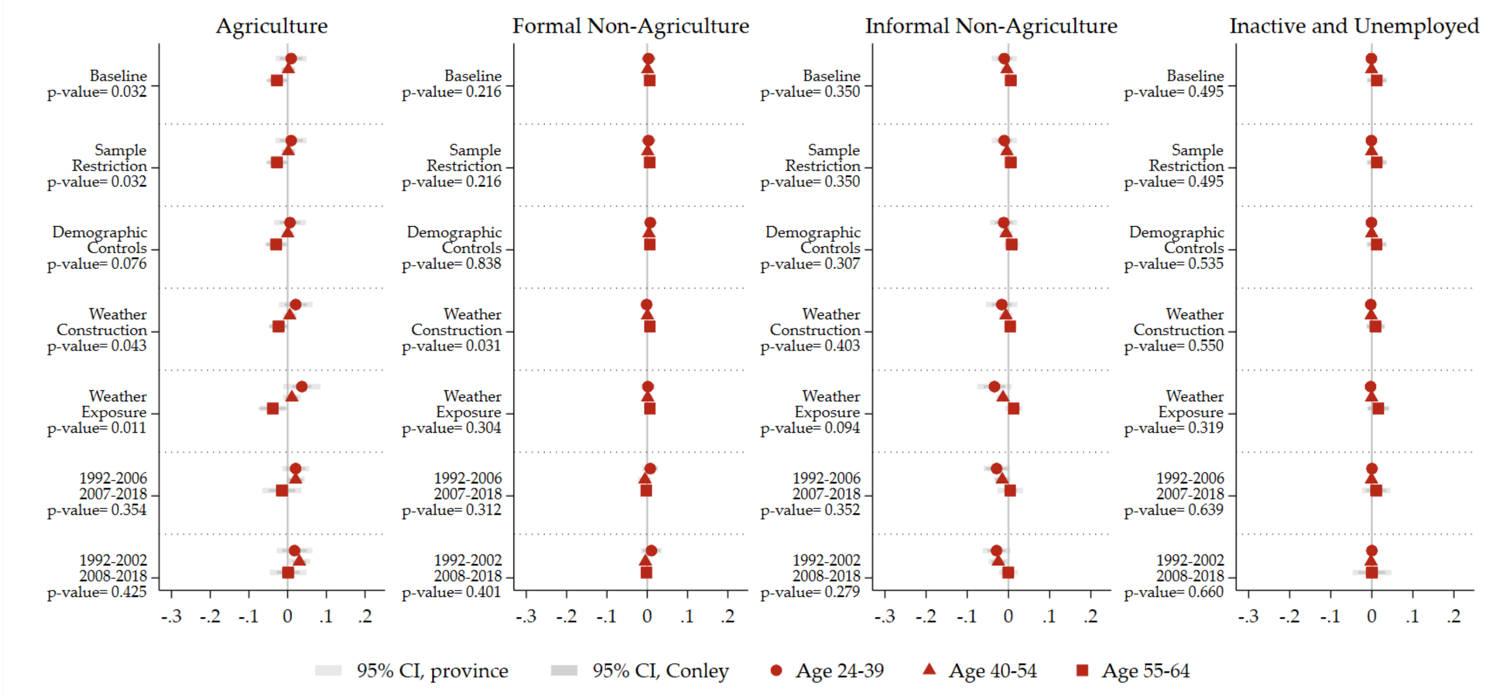
Figure C11: Wet-bulb Temperature and Sectoral Labor Share by Age Group



Notes: Results from estimating equation (13). Other notes are similar to Figure C3. p-values from the test of significant age cohort differences using standard errors clustered at the province level are reported. The results are qualitatively similar when using Conley standard errors.

Figure C12: Wet-bulb Temperature and Sectoral Labor Share by Age Group

Results from Long Differences Approach with Degree Days Measures (DD9)



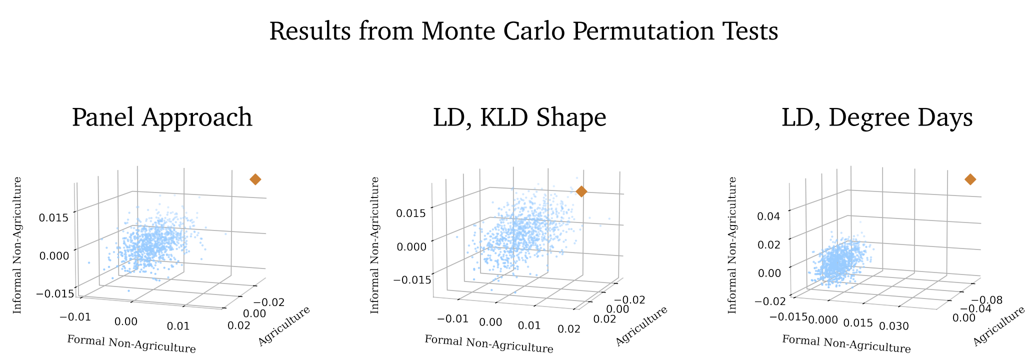
Notes: Results from estimating equation (13). Other notes are similar to Figure C3. p-values from the test of significant age cohort differences using standard errors clustered at the province level are reported. The results are qualitatively similar when using Conley standard errors.

C2 Placebo Test

As an additional check on the econometric specifications, I conduct a placebo test with Monte Carlo analyses of equations (11) and (13) 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 (11) and (13). 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 C13 presents the joint distribution of the estimated coefficients with random assignment for agricultural, formal, and informal non-agricultural employment share outcomes. 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.

Figure C13: Placebo Test: Wet-bulb Temperature and Sectoral Labor Share



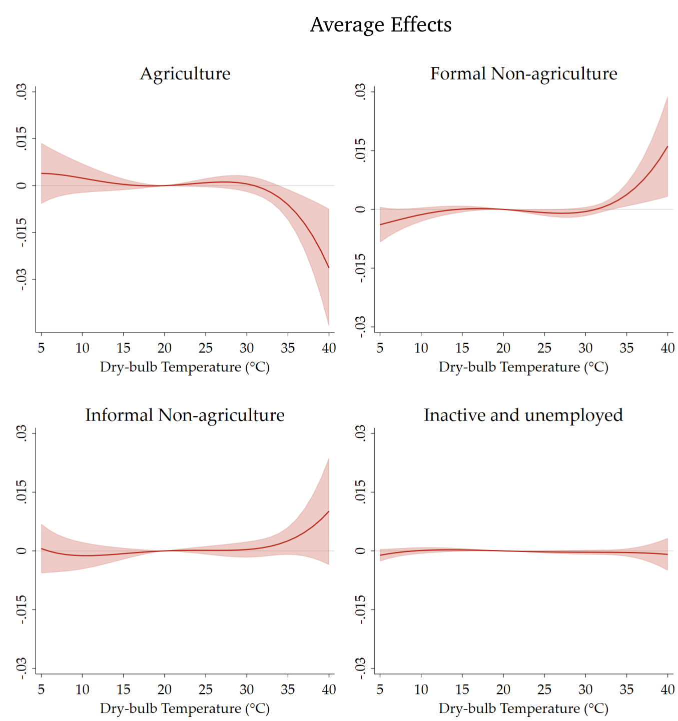
Notes: Each graph presents the (joint) distribution of estimated coefficients (circles) of temperature effects on sectoral employment shares, which are obtained from 1,000 Monte Carlo simulations where the weather series of one analysis unit is randomly reassigned to another unit's employment share series. Diamonds represent the baseline estimates from the panel approach and the long differences approach using KLD measures (shape) degree days measures (DD27).

C3 Additional Results using Dry-Bulb Temperatures

I estimate equation (11) using fourth-order polynomials of dry-bulb temperatures. As seen in Appendix Figure C14, similar to the results using wet-bulb temperatures, hot dry-bulb temperatures are associated with a decrease in agricultural labor share and increases in formal and informal non-agricultural employment shares, but do not affect the share of unemployed and inactive workers. If anything, the temperature effect on informal non-agricultural labor share is somewhat less precisely estimated. Note that because wet-bulb temperatures are always lower than dry-bulb temperatures, the dry-bulb temperature cutoff above which labor reallocation effects are concentrated is higher, at approximately 30°C. This finding is consistent with Hsiang (2010), who shows that above 29°C surface temperature, the production-temperature response steepens for all industries and services in the Caribbean and Central America regions.

Figure C14: Dry-bulb Temperature and Sectoral Labor Share

Results from Panel Approach with Fourth-Order Polynomials



Notes: This figure represents a predicted sectoral employment share-temperature response function, estimated with equation (11), where estimates are from a fourth-order polynomial in daily average dry-bulb temperature. Shaded areas are 95% confidence interval. Robust standard errors are clustered at the province level.

D Additional Tables and Figures

D1 Additional Tables

Table D1: Variation in Temperature Variables under Different Sets of Fixed Effects

Panel A: Degree Days in Panel Approach

	DD27			DD9		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\sigma\_e(°C)$	$\|e\| > 0.5°C(\%)$	$\|e\| > 1°C(\%)$	$\sigma\_e(°C)$	$\|e\| > 0.5°C(\%)$	$\|e\| > 1°C(\%)$
No FE	7.1	97.9	97.0	17.3	99.0	98.1
Province-age group FE	2.9	24.5	19.1	8.7	41.7	39.8
Province-age group + region-year FE	1.3	31.0	20.0	4.5	40.0	33.6
Province-age group + region-age group-year FE	1.3	31.4	20.2	4.5	39.8	33.5

Panel B: Degree Days in Long Differences Approach

	DD27			DD9		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\sigma\_e(°C)$	$\|e\| > 0.5°C(\%)$	$\|e\| > 1°C(\%)$	$\sigma\_e(°C)$	$\|e\| > 0.5°C(\%)$	$\|e\| > 1°C(\%)$
No FE	3.6	96.8	92.3	2.4	91.0	82.1
Region-age group FE	1.0	32.1	16.0	1.8	33.3	24.4

Panel C: KLD Measures in Long Differences Approach

	KLD Shape	KLD Location
	(1)	(2)
	σ_e	σ_e
No FE	0.09	0.20
Region-age group FE	0.05	0.08

Notes: This table summarizes regressions of degree days (Panels A and B) and of KLD measures (Panel C) on different set of fixed effects and how much of the variation remain. In Panels A and B, columns (1) and (3) report the standard deviation of the residuals (remaining temperature variation) in degree Celsius, and columns (2/3) and (5/6) report the fraction of residuals with an absolute value greater than 0.5/1°C over the reference period.

Table D2: Wet-bulb Temperature and Labor Shares by Industry

Results from Panel Approach with Degree Days

	Agriculture	Non-Agriculture		
	(1)	(2) Less Skill-Intensive	(3) Medium Skill-Intensive	(4) High Skill-Intensive
DD9	-0.0020	0.0039	0.0014	-0.0032
	(0.0082)	(0.0072)	(0.0013)	(0.0013)
.	[0.0064]	[0.0057]	[0.0010]	[0.0018]
DD27	-0.0466	0.0351	0.0130	0.0063
	(0.0126)	(0.0079)	(0.0023)	(0.0040)
.	[0.0155]	[0.0132]	[0.0024]	[0.0030]
Mean Outcome	0.45	0.33	0.03	0.09
Adjusted R2	0.82	0.64	0.47	0.46
Observations	1707	1707	1707	1707
Province × Age Group FE	x	x	x	x
Region × Age Group × Year FE	x	x	x	x

Notes: This table presents the effect of 1 SD increase in degree days above 27°C or below 9°C on the share of employment in each industry group. Unit of analysis is province-year. Less skill-intensive industries include those classified as low-tech manufacturing, less knowledge-intensive services, construction and mining; medium skill-intensive industries include those classified as medium-tech manufacturing and public utilities; high skill-intensive industries include those classified as high-tech manufacturing and knowledge-intensive services according to the Statistical Classification of Economic Activities in the European Community. The mean educational attainment of workers in each industry is presented in Appendix Table B3. 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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table D3: Decomposing Changes in Sectoral Labor Share Shares: 1992-2018

	Agriculture	Informal Non-Agriculture	Formal Non-Agriculture
	(1)	(2)	(3)
Total Change	-0.314	0.111	0.203
Within Province	-0.312	0.121	0.200
Between Province	-0.002	-0.009	0.002

Notes: This table presents the results from the decomposition exercise following McCaig and Pavcnik (2018). I decompose the change in the share of workers in each sector in total employment between 1992 and 2018, denoted by ΔS, into within and between province shifts, respectively: $\Delta S_t = S_t - S_{t-1} = \sum_p \Delta s_{pt} h_p + \sum_p \Delta h_{pt} s_p$ where h_{pt} is the share of province p's employment in total employment at time t, s_{pt} is the share of workers in sector s in total employment in province p, $h_p = 0.5(h_{pt} + h_{pt-1})$, and $s_p = 0.5(s_{pt} + s_{pt-1})$. The first summation term captures the importance of mobility of workers across sectors within a province, and the second summation captures the prevalence of mobility of workers across provinces as sources of changes in aggregate sectoral employment shares. Estimates based on VLSS 1992/1993 and VHLSS 2018. Sample includes workers aged 24-64 inclusive. Survey sampling weights included. The results are qualitatively similar when using other household survey rounds as the start point.

Table D4: Wet-bulb Temperature and Migration Responses

	In Migration		Out Migration		Net Migration	
	(1)	(2)	(3)	(4)	(5)	(6)
DD9	0.0083	-0.0020	0.0144	-0.0324	-0.0072	0.0305
	(0.0078)	(0.0172)	(0.0278)	(0.0613)	(0.0319)	(0.0557)
.	[0.0035]	[0.0140]	[0.0304]	[0.0429]	[0.0318]	[0.0368]
DD27	0.0614	0.0622	-0.0811	-0.0336	0.1445	0.0975
	(0.0535)	(0.0565)	(0.0752)	(0.0981)	(0.0913)	(0.1194)
.	[0.0420]	[0.0466]	[0.0735]	[0.0851]	[0.0820]	[0.1022]
N	312	312	312	312	312	312
Mean Outcome	5.060	5.060	6.712	6.712	-1.653	-1.653
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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table D5: Hot Wet-bulb Temperatures and Primary Sectoral Labor Share, by Gender and Education

	Agriculture	Formal Non-Agriculture	Informal Non-Agriculture	No Employment
	(1)	(2)	(3)	(4)
Male - Female	-0.0002	-0.0014	0.0010	0.0006
	(0.0019)	(0.0012)	(0.0019)	(0.0006)
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	0.0033	0.0007	-0.0052	0.0013
	(0.0030)	(0.0016)	(0.0022)	(0.0009)
Mean Outcome	0.4501	0.1710	0.2769	0.0865
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.

Table D6: Wet-bulb Temperature and Rice Price

	Price Harvest				Price Sold			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mean	Mean	Median	Median	Mean	Mean	Median	Median
DD9	-0.0043	-0.0055	-0.0032	-0.0044	-0.0022	-0.0028	0.0001	-0.0006
	(0.0038)	(0.0045)	(0.0040)	(0.0049)	(0.0039)	(0.0045)	(0.0038)	(0.0048)
.	[0.0029]	[0.0030]	[0.0031]	[0.0032]	[0.0029]	[0.0027]	[0.0031]	[0.0031]
DD27	-0.0024	-0.0032	-0.0060	-0.0046	-0.0033	0.0048	-0.0088	-0.0038
	(0.0072)	(0.0080)	(0.0093)	(0.0090)	(0.0113)	(0.0078)	(0.0160)	(0.0126)
.	[0.0077]	[0.0084]	[0.0087]	[0.0096]	[0.0088]	[0.0093]	[0.0122]	[0.0127]
N	415	415	415	415	415	415	415	415
Mean Outcome	4.13	4.13	4.12	4.12	4.19	4.19	4.15	4.15
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. In Columns (1)-(4), dependent variables are price of rice at harvest, measured as the province-level mean/median of household-level revenue at harvest divided by output (thousand VND/kg). In Columns (5)-(8), dependent variables are price of rice sold, measured as the province-level mean/median of household-level revenue sold divided by output sold (thousand VND/kg). 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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids. All regressions are weighted by household survey sampling weights interacted with cultivated areas.

Table D7: Wet-bulb Temperature and Household Consumption by Trade Openness

Trade openness is proxied by distance to the nearest major seaport

	Total Consumption		Food Consumption		Nonfood Consumption		Food Share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DD27 × (Open=0) (N)	-0.0087	-0.0098	-0.0015	-0.0019	-0.0071	-0.0080	0.0011	0.0014
	(0.0041)	(0.0035)	(0.0014)	(0.0012)	(0.0033)	(0.0027)	(0.0014)	(0.0011)
.	[0.0036]	[0.0036]	[0.0015]	[0.0015]	[0.0024]	[0.0025]	[0.0010]	[0.0010]
DD27 × (Open=1) (T)	0.0910	0.0775	0.0137	0.0123	0.0773	0.0652	-0.0189	-0.0115
	(0.0543)	(0.0717)	(0.0100)	(0.0141)	(0.0458)	(0.0584)	(0.0035)	(0.0068)
.	[0.0626]	[0.0719]	[0.0125]	[0.0136]	[0.0512]	[0.0595]	[0.0080]	[0.0083]
p-value (N) = (T)	0.0730	0.2272	0.1378	0.3196	0.0715	0.2141	0.0000	0.0679
Mean Outcome	1.1346	1.1346	0.4402	0.4402	0.6944	0.6944	0.4809	0.4809
Observations	569	569	569	569	569	569	569	569
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: Each panel presents the effect of 1 SD increase in degree days above 27°C, separately for tradable and non-tradable markets, on per capita monthly household consumption (2010 million VND). "Open" is an indicator that takes value 1 if the distance from a province centroid to the nearest major port is below the 70th percentile (approximately 200 km) and 0 otherwise. All regressions control for other weather variables (cold temperatures, second-order polynomials of precipitation and wind speed) and their interactions with the 'Open' dummy. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 150 km and and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table D8: Wet-bulb Temperature and Crop Yields

	Rice (All)		Rice (Winter-Spring)		Maize		Grains	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DD9	-0.0009	-0.0009	-0.0013	-0.0013	0.0016	0.0012	-0.0011	-0.0009
	(0.0016)	(0.0014)	(0.0012)	(0.0011)	(0.0024)	(0.0025)	(0.0010)	(0.0010)
.	[0.0021]	[0.0013]	[0.0027]	[0.0009]	[0.0018]	[0.0016]	[0.0015]	[0.0010]
DD27	-0.0198	-0.0174	-0.0404	-0.0346	-0.0088	0.0049	-0.0193	-0.0157
	(0.0078)	(0.0071)	(0.0131)	(0.0118)	(0.0101)	(0.0074)	(0.0076)	(0.0068)
.	[0.0084]	[0.0066]	[0.0144]	[0.0101]	[0.0082]	[0.0059]	[0.0082]	[0.0063]
N	520	520	520	520	515	515	520	520
Mean Outcome	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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table D9: Wet-bulb Temperature and Agricultural Planting Area

	Rice (All)		Rice (Winter-Spring)		Maize		Grains	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DD9	-0.0042	0.0006	0.0073	0.0073	-0.1146	-0.0890	-0.1088	-0.0783
	(0.0392)	(0.0266)	(0.0169)	(0.0153)	(0.1017)	(0.0826)	(0.1091)	(0.0839)
.	[0.0390]	[0.0157]	[0.0149]	[0.0105]	[0.0823]	[0.0508]	[0.0991]	[0.0554]
DD27	-0.4749	-0.2502	0.0213	0.0186	0.2712	0.1039	-0.1736	-0.1165
	(0.3850)	(0.1513)	(0.1028)	(0.0338)	(0.1700)	(0.1138)	(0.4050)	(0.1847)
.	[0.3753]	[0.1395]	[0.0723]	[0.0478]	[0.2437]	[0.1176]	[0.4882]	[0.1792]
N	520	520	520	520	520	520	520	520
Mean Outcome	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 150 km and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

Table D10: Wet-bulb Temperature and Firm-Level Labor Productivity

	Mining and Quarrying		Construction	
	(1)	(2)	(3)	(4)
DD27	-0.0040	-0.0074	-0.0014	0.0000
	(0.0049)	(0.0057)	(0.0030)	(0.0029)
DD27 × (Firm Size < 30 workers)	-0.0120		-0.0008	
	(0.0036)		(0.0019)	
DD27 × (Firm Age >= 10 years old)		-0.0087		-0.0061
		(0.0028)		(0.0010)
Total Temperature Effects	-0.0160	-0.0161	-0.0022	-0.0061
	(0.0056)	(0.0056)	(0.0029)	(0.0030)
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
Observations	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 < 9°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 D11: Wet-bulb Temperature and Sectoral Labor Share by Long-run Temperatures

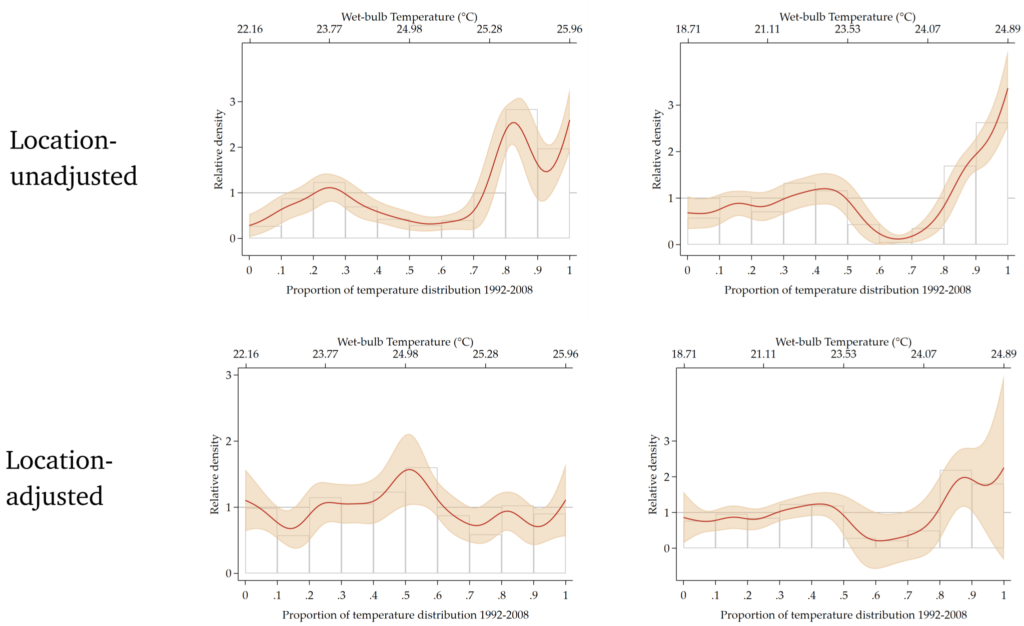
Results from Panel Approach

	(1) Agriculture	(2) Formal Non-Agriculture	(3) Informal Non-Agriculture	(4) Inactive and Unemployed
DD27 × Hot (H)	-0.0103	0.0027	0.0080	-0.0005
	(0.0014)	(0.0010)	(0.0015)	(0.0006)
.	[0.0016]	[0.0010]	[0.0004]	[0.0006]
DD27 × Less Hot (C)	-0.0055	0.0030	0.0031	-0.0006
	(0.0018)	(0.0010)	(0.0015)	(0.0008)
.	[0.0022]	[0.0009]	[0.0018]	[0.0006]
p-value (H) = (C)	0.0001	0.5690	0.0001	0.8665
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. Dependent variables are shares of employment in each sector. 'Hot' ('Less Hot') is an indicator that takes value 1 if the long-term mean temperatures of the province is above (below) the country's median level and zero otherwise. All regressions control for weather variables and their interactions with 'Hot' dummy. Robust standard errors clustered at the province level are in parentheses. Conley standard errors that allow for spatial correlation up to 150 km and and serial correlation up to five lags are in brackets. Province distances are computed from province geographic centroids.

D2 Additional Figures

Figure D1: Change in wet-bulb temperature distribution in two provinces: 1992-2008 vs. 2009-2018

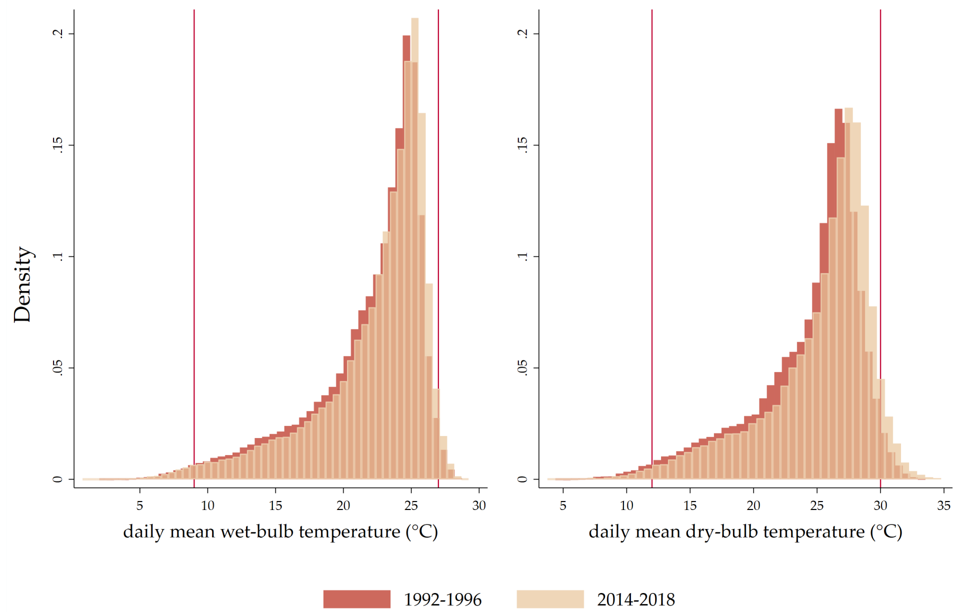


(a) Province A	(b) Province B
overall divergence = 0.34	overall divergence = 0.35
location = 0.25, shape = 0.08	location = 0.09, shape = 0.25
Δmean temperature = 0.39°C	Δmean temperature = 0.40°C

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 Province A experiences a relatively smooth righward shift in the whole temperature distribution (location effect), Province B observes a polarization of temperature distribution with fewer mild days and more hot days in recent years (shape effect).

Figure D2: Distribution of daily mean temperatures 1992-2018

Panel A: Period 1992-2018



Panel B: Period 1992-1996 vs. 2014-2018

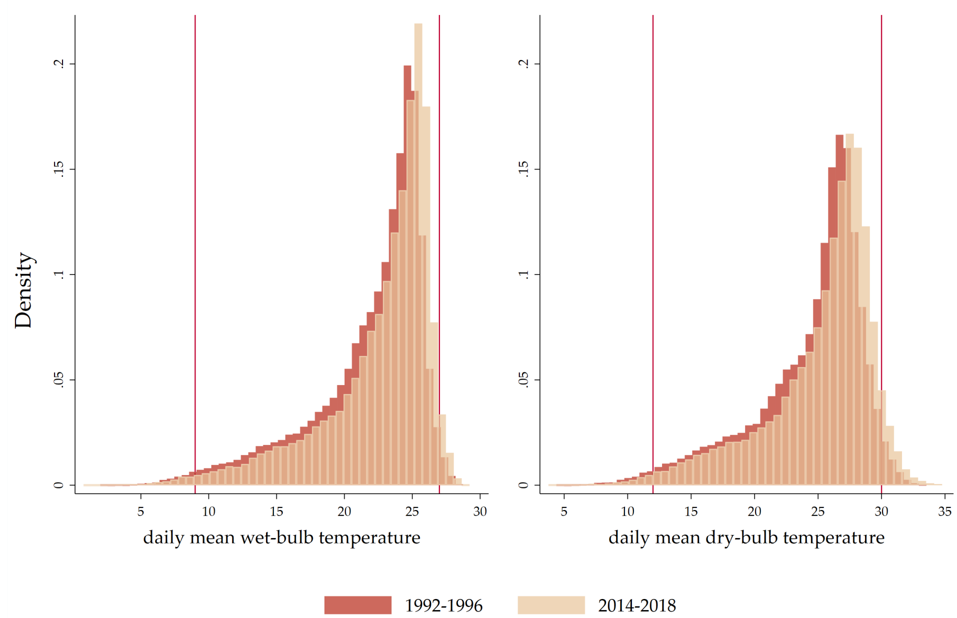


Figure D3: Interview Month of an Average Person in each Province: 1992-2018

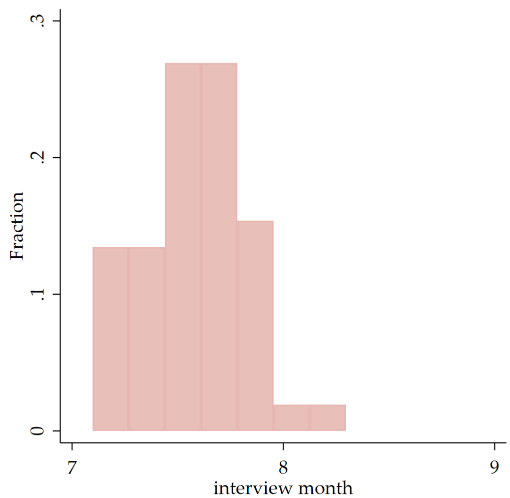
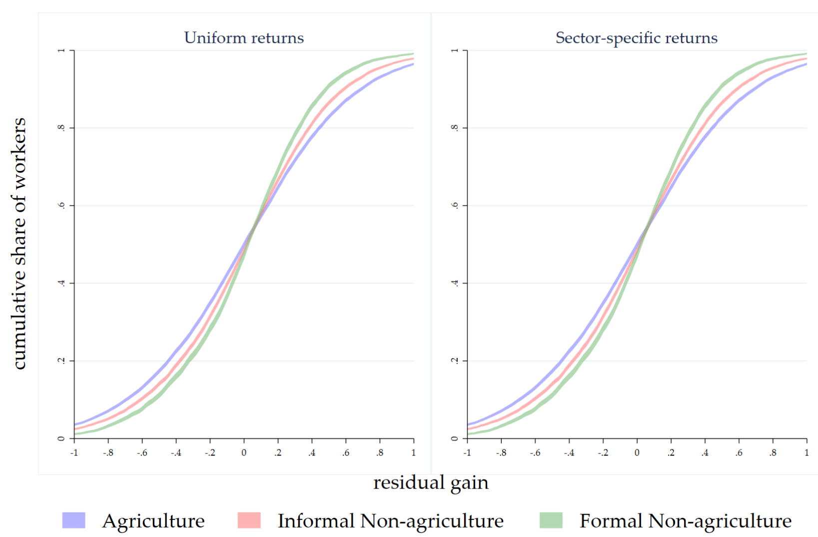
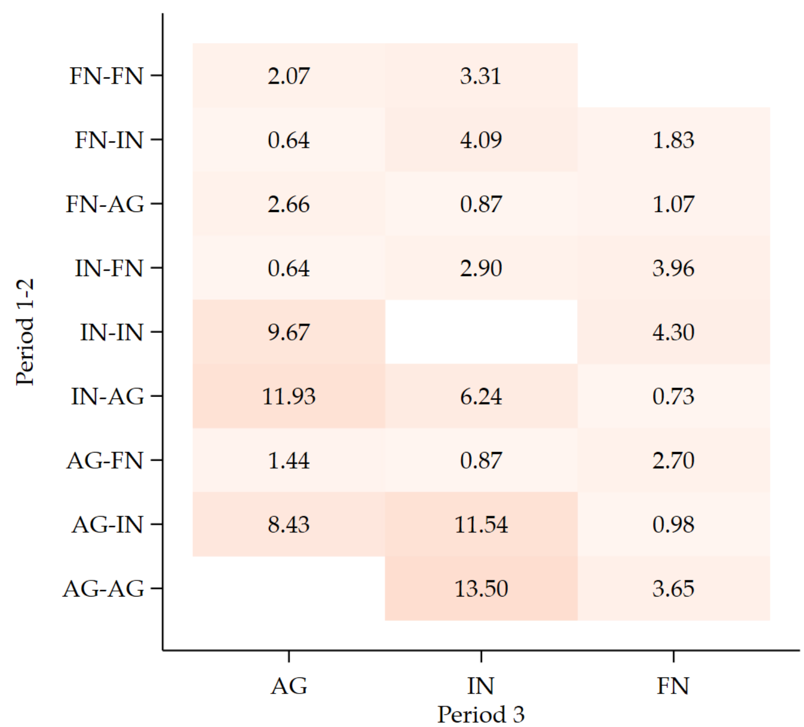


Figure D4: First-order stochastic dominance among sector-specific residualized distributions



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.

Figure D5: Share of Switchers



Notes: This figure presents the percentage share (%) of switchers who primarily worked in one of sectors listed on the y-axis in the first and second periods, and on the x-axis in the third period, among all workers who switched at least once, using data from five VHLSS three-wave individual panels 2002-2004-2006 to 2014-2016-2018. "AG" denotes agriculture, "IN" denotes informal non-agriculture, and "FN" denoted formal non-agriculture. Among switchers who worked in agriculture in the first period, and in informal non-agriculture in the second period, only 4.7% was able to work a formal non-agricultural job in the third period (0.98/(8.43+11.54+0.98)).

E 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:

$$Y_{jt} = A_{jt}(\text{WBT})\Pi_k(X^k_{jt}(\text{WBT}))^{\theta_k} \tag{S10}$$

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 or labor effort per unit time worked (Somanathan et al. 2021; Graff Zivin and Neidell 2014).

To measure marginal product of labor, one can take natural logs of equation (S10) and obtain the empirical model:

$$y_{jt} = \theta_0 + \sum_k \theta_k x^k_{jt} + u_{jt} \tag{S11}$$

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. θ_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 (S11) 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 (S11) 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:

$$MP_{Ljt} = \hat{\theta}_l \frac{y_{jt}}{l_{jt}} \tag{S12}$$

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

$$MP_{Ljt} = f(\mathrm{WBT}_{pt}) + g(R_{pt}) + \gamma_p + \gamma_{rt} + \varepsilon_{prt} \tag{S13}$$

Similar to the main analysis, $f(\mathrm{WBT}_{pt})$ can be represented by cumulative temperature bins, degree day bins, and fourth-order polynomials. In the most parsimonious model where $f(\mathrm{WBT})$ is a piece-wise linear function:

$$f(\mathrm{WBT}_{pt}) = \begin{cases} \sum_{d=1}^{365} \beta_9 (9 - \mathrm{WBT}_{dpt}) & \text{if} \quad 0 \leq \mathrm{WBT} < 9 \\ 0 & \text{if} \quad 9 \leq \mathrm{WBT} < 27 \\ \sum_{d=1}^{365} \beta_{27} (\mathrm{WBT}_{dpt} - 27) & \text{if} \quad \mathrm{WBT} \geq 27 \end{cases} \tag{S14}$$

The estimated coefficients $\hat{\beta}_9$ and $\hat{\beta}_{27}$ represent the effects of one additional degree day below 9°C or above 27°C, respectively, on sectoral marginal product of labor during the fiscal year reference period.

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.

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.

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.

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*, 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.