

# Who Benefits from Place-Based Industrial Policies? Labor Market Adjustments and Household Welfare in Vietnam

Trinh Pham\*

This Version: April 2026

## Abstract

Governments routinely evaluate economic policies by measuring their effects on formal employment and registered firms. But in settings where most workers earn their livelihoods outside the formal sector, this approach may understate total benefits and mischaracterize who gains. This paper examines this question in the context of industrial zones in Vietnam. To guide the empirical analysis, I develop a model of segmented local labor markets in which formal employment expands but remains constrained for less-educated workers. The model predicts that higher formal earnings raise demand for locally produced non-tradable goods and services, increasing returns to informal activities through an indirect local demand channel. Using nationally representative household survey data from 2002 to 2020 and a staggered difference-in-differences design, I find strong support for these predictions: less-educated households reallocate labor from agriculture toward informal non-farm sector, increase labor diversification, and raise income. Informal income accounts for three-quarters of their non-farm income gains, driven by non-tradable business activities. These gains translate into higher consumption and increased investment in children's schooling. Assessments based solely on formal employment would capture only a quarter of the income gains for less-educated households, substantially understating the welfare impact of industrial zones.

**Keywords:** place-based policy, labor reallocation and informality, household welfare

**JEL Codes:** J46, O17, R11

---

\*KDI School of Public Policy and Management. Email: tpham@kdischool.ac.kr; 263 Namsejong-ro, Sejong-si, Republic of Korea. This project is supported by funding from KDI School. I would like to thank participants at NEUDC 2025, World Bank-KDI School 6th Development Impact Conference, KDI School, and National Cheng Kung University for helpful suggestions on previous versions of the paper. I also thank Dung Dang and Hien Truong for helpful research assistance. Any remaining errors are mine. A previous version of the paper was circulated as "Who Benefits from Industrialization? Labor Market Adjustments and Household Welfare in Vietnam."

# 1 INTRODUCTION

Governments routinely evaluate economic policies by measuring their effects on formal employment, registered firms, and taxable income. In much of the world, however, the majority of workers and households earn their livelihoods partly or entirely outside the formal sector (Bonnet et al., 2019). If policies generate substantial effects through informal channels, then standard evaluation approaches may understate total benefits and mischaracterize who gains. This paper provides evidence that this concern is empirically large in the context of place-based industrial zones. These programs are among the most widely adopted development policy instruments, with more than 5,400 zones operating across low- and middle-income countries by 2019 (UNCTAD, 2019). Yet the evidence on whether these programs improve welfare in surrounding communities is mixed, with large effects documented in some settings and modest or null effects in others. This paper suggests that part of this apparent heterogeneity reflects what evaluations measure rather than whether policies work.

To guide the analysis, I develop a simple conceptual framework for a local economy with three production activities (agriculture, formal non-agriculture, and informal non-agriculture) and two household types that differ in their access to formal employment. Industrial zone establishment expands the scale and productivity of the formal sector, raising formal wages and employment opportunities, though access for less-educated workers remains constrained. Higher formal earnings increase local demand for non-tradable goods and services, which are produced and consumed locally through informal activities. This raises the equilibrium return to informal non-farm work, inducing less-educated households to reallocate labor away from agriculture and toward informal non-farm sector. Less-educated households can therefore benefit through two margins: limited entry into formal employment as access expands, and reallocation of non-formal labor toward informal activities whose returns rise with local demand.

Importantly, this adjustment needs not take the form of a clean transition from one sector to another. When formal employment is constrained and informal opportunities are expanding, less-educated households may spread labor across multiple activities, temporarily diversifying their labor portfolios across agriculture, formal, and informal work. This within-household diversification is a policy-relevant margin of adjustment: if it is large enough, individual-level employ-

ment patterns that appear to favor educated workers may mask more inclusive welfare gains at the household level. These mechanisms yield four testable predictions: (i) formal employment gains should be larger for educated workers; (ii) less-educated households should reallocate labor from agriculture into informal non-farm activities; (iii) less-educated households are expected to temporarily increase labor diversification across sectors; and (iv) informal non-farm income may contribute significantly to total income gains for less-educated households, even when formal employment gains are modest for this group.

I test these predictions using Vietnam's large-scale expansion of industrial zones over the last two decades. Vietnam shares key structural features with many low- and middle-income countries: a large informal sector, widespread household engagement in multiple economic activities, and sharp educational stratification in access to formal employment. The country's industrial zone program also offers favorable conditions for credible identification. The program expanded rapidly from fewer than 70 zones in 2000 to over 300 by 2020, with establishment staggered across districts in a way that provides substantial temporal and geographic variation. I exploit this variation and estimate dynamic treatment effects using the difference-in-differences event-study estimator of de Chaisemartin et al. (2024), which is designed to accommodate staggered adoption and heterogeneous effects. I define a district as treated if any part of the district lies within 15 kilometers of an industrial zone, a threshold supported by spatial ring analysis showing that impacts attenuate sharply beyond this distance. Combining this treatment measure with ten waves of nationally representative Vietnam Household Living Standards Survey (VHLSS) data from 2002 to 2020, the design identifies how localized industrialization affects labor market outcomes and household economic behavior.

The analysis yields four main findings, each corresponding to a prediction of the framework. First, industrial zones induce substantial sectoral reallocation, but access to formal employment remains sharply stratified by education. Workers with at least a high school diploma shift primarily into formal non-agricultural employment, while less-educated workers reallocate out of agriculture predominantly into informal non-agricultural activities. The expansion in informal employment reflects entry as a primary livelihood rather than a supplemental activity: significant effects appear in both the primary-job and sole-job measures. Second, households without a high school graduate exhibit a pronounced increase in labor diversification across sectors: the response oper-

ates through both members specializing in different sectors and individuals combining jobs across sectors, peaks around eight to ten years after zone establishment, and then gradually attenuates. Households with educated members show no comparable within-member response, and households where all members hold a high school diploma show no diversification response on any margin, consistent with the model's prediction that diversification is a compensating strategy for households with constrained formal access. Third, despite limited formal job access, less-educated households experience significant income gains driven primarily by higher profits from informal household enterprises concentrated in non-tradable services such as food preparation and accommodation. Informal non-farm income accounts for approximately 75% of total non-farm income gains for this group, with formal income contributing the remaining 25%. The concentration of gains in non-tradable rather than tradable business activity is consistent with the local demand channel, rather than a general improvement in the business environment, driving the informal income response. Fourth, these income gains are accompanied by higher non-food consumption, increased health and education spending, and higher school enrollment among children of less-educated households, suggesting that the informal income channel translates into tangible improvements in living standards and human capital investment.

Two additional pieces of evidence speak to how the local demand channel operates. Industrial zone establishment selectively attracts non-local workers into formal employment, with a 62% increase in the share of non-registered formal workers relative to the pre-treatment mean and no corresponding inflow of agricultural workers. This migration response is part of the mechanism through which zones generate local spending, but the welfare gains accrue to incumbent households: restricting the sample to long-term registered residents yields near-identical estimates. Furthermore, the diversification and non-tradable business responses are substantially larger in districts with higher baseline skill intensity, where formal income gains and thus local demand shifts are larger, a pattern more consistent with demand-driven spillovers than with risk diversification. These findings are robust to alternative estimators, inference procedures, sample restrictions, and spatial and sampling checks, including a conventional two-way fixed-effects (TWFE) specification that yields qualitatively consistent but attenuated estimates, a donut exclusion test suggesting the baseline is conservative, and Monte Carlo permutation tests confirming that all results fall well outside the distribution of placebo estimates.

This paper makes three contributions. First, it provides evidence that the gap between formal-sector evaluation and actual welfare effects can be quantitatively large in economies where informality is prevalent. This concern is likely relevant across a wide range of policy settings, and this paper explores it in the context of place-based industrial programs, which are among the most widely adopted development policy instruments and are expanding rapidly in high-informality economies (UNCTAD, 2019). Evidence from advanced economies has shown that spatially targeted programs can generate persistent local employment gains (Garin & Rothbaum, 2025; Kline & Moretti, 2014), that tradable-sector expansion creates non-tradable employment through local demand (Moretti, 2010), and that local productivity growth can disproportionately benefit less-educated workers (Hornbeck & Moretti, 2024).<sup>1</sup> These findings are derived predominantly from settings where economic activity is formally registered and labor tends to move relatively freely across sectors.

In developing economies, however, substantial frictions limit worker reallocation (Donovan et al., 2023; Fields, 2011) and a large share of the workforce operates in informal household enterprises (La Porta & Shleifer, 2014; Maloney, 2004; Ulyssea, 2020). Much of the existing evidence examines how informality mediates the effects of adverse economic shocks: structural work on trade liberalization shows that the informal sector mitigates the unemployment consequences of negative labor demand shocks but can amplify misallocation and real income losses (Dix-Carneiro et al., 2026). This paper provides evidence of a complementary role: when formal-sector expansion generates local demand, the informal sector can serve as the primary channel through which welfare gains reach less-educated households. I show that in Vietnam, formal employment and wage gains account for only one quarter of total non-farm income gains for households with limited formal access; the remaining three quarters accrue through informal non-tradable enterprise activity. These findings are rationalized by a framework that draws on models of segmented labor markets (Fields, 2011; Harris & Todaro, 1970), local demand multipliers (Moretti, 2010), and household labor allocation under constrained access to formal employment. The framework identifies two conditions under which place-based industrial expansion generates broad-based welfare gains: zones must successfully create formal activity, and the local economy must have an informal sector flexible enough to absorb rising demand.

---

<sup>1</sup>See Neumark and Simpson (2015) for a comprehensive review of the place-based policy literature.

Second, the paper contributes to a growing empirical literature on the effects of special economic zones in developing countries. Prior work has documented that zones increase firm entry, employment, and productivity in China (Y. Lu et al., 2019; Wang, 2013), affect local human capital investment (F. Lu et al., 2023), induce intra-household resource reallocation (Zhao & Qu, 2024), promote structural change from agriculture to manufacturing and services in India and Vietnam (Gallé et al., 2024; Tafese et al., 2025), and improve household wealth near zones across African countries (Abagna et al., 2025). In Indonesia, by contrast, zones appear to have had limited effects on local economies (Rothenberg et al., 2025). This paper extends this body of evidence by documenting each stage of the transmission from zone establishment to welfare gains for less-educated households: formal job creation, selective in-migration of formal workers, expansion of non-tradable informal businesses, and income gains for incumbent households with limited formal access. The concentration of gains in non-tradable rather than tradable activities, and the substantially stronger responses in more skill-intensive districts, are consistent with the local demand channel emphasized in the framework.

Third, the paper documents a margin of household adjustment to new employment opportunities that has received limited attention in the literature: temporary cross-sector labor diversification among skill-constrained households. A large literature has documented multi-activity livelihoods in developing economies, often interpreted as risk management under incomplete markets (Barrett et al., 2001; Dercon, 2002; Rosenzweig & Stark, 1989) or as a consequence of limited access to stable formal employment (Fields, 2011; Harris & Todaro, 1970; Maloney, 2004). I find that less-educated households respond to zone establishment by temporarily spreading labor across agriculture, formal, and informal work, with diversification peaking around eight to ten years before attenuating. The response traces a gradient across household types: strong diversification with no educated members, across-member sorting only in mixed-skill households, and no response when all members hold a high school diploma. This gradient appears more consistent with demand-driven opportunity expansion than with risk diversification, as the response is substantially larger in more skill-intensive districts where formal income gains and local demand shifts are likely greater, though this heterogeneity should be interpreted as suggestive.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting, documents the expansion of manufacturing activity following zone establishment, and character-

izes the labor market context. Section 3 develops a conceptual framework that generates testable predictions about sectoral reallocation, household diversification, and welfare. Section 4 describes the data and empirical strategy. Section 5 presents the main empirical results. Section 6 reports robustness checks. Section 7 discusses the role of migration and the broader implications for evaluating place-based policies. Section 8 concludes.

## 2 INSTITUTIONAL BACKGROUND AND LABOR MARKET CONTEXT

### 2.1 Industrial Zones in Vietnam

In Vietnam, industrial zones are designated areas intended to concentrate industrial production, manufacturing, and related services. Their core objectives are to attract foreign direct investment, promote export-oriented growth, and create employment opportunities. To support these goals, zones are equipped with dedicated infrastructure and offer a variety of investment incentives, making them appealing to both domestic and international firms.

The planning and development of these zones involve close coordination between central and provincial authorities, aligning with Vietnam's broader socio-economic strategies.<sup>2</sup> At the national level, the Ministry of Planning and Investment oversees the planning, regulation, and promotion of industrial zones, ensuring new zones are integrated into national and provincial master plans and formally approved by the Prime Minister. Provincial People's Committees play a central role in identifying potential locations, managing local implementation, and determining region-specific incentives and infrastructure investments. Once a zone is approved, provincial management boards are responsible for daily operations, administrative services, and monitoring compliance with investment and environmental regulations.

Firms locating in industrial zones benefit from a package of fiscal and administrative incentives designed to lower the cost of investment. On the tax side, the Corporate Income Tax Law (No. 14/2008/QH12) and subsequent implementing decrees provide firms in industrial zones with a two-year exemption from corporate income tax (CIT) from the first year of profitability, followed by a 50% reduction for the next four years, with the standard CIT rate of 20% apply-

---

<sup>2</sup>See Decree 35/2022/NĐ-CP on Regulations on the Management of Industrial Zones and Economic Zones for more details. <https://vanban.chinhphu.vn/?pageid=27160&docid=205861>.

ing thereafter.<sup>3</sup> Beyond corporate tax, firms can import machinery and equipment not produced domestically duty-free if these inputs are used to form fixed assets.<sup>4</sup> On the administrative side, zone management boards act as a single point of contact for investment licensing, construction permits, and regulatory approvals, allowing firms to obtain all necessary authorizations through one office rather than coordinating across multiple government agencies. Firms also benefit from zone-provided infrastructure, including roads, electricity, water supply, and wastewater treatment, which would otherwise require costly private investment, as well as preferential land rental rates with partial or full exemptions depending on zone location and project characteristics.

Figure 1 illustrates the growth trajectory of industrial zones in Vietnam from 1990 to 2020. By the mid-1990s, industrial zone development was limited, with fewer than 20 zones nationwide. The issuance of Decree 192/CP in late 1994 introduced the country's first comprehensive legal framework for industrial parks, marking the beginning of more structured development. The late 1990s saw a steady increase in the number of zones, with rapid growth occurring in the early 2000s following reforms to the Foreign Direct Investment Law. These reforms liberalized investment conditions, triggering a surge in zone establishment, with the total number rising from around 70 to over 200 within a few years. Continued growth through the late 2000s and 2010s coincided with Decree 29/2008, which simplified industrial zone management and administrative procedures. By 2020, over 400 industrial zones had been planned, with approximately 300 officially established and operational.

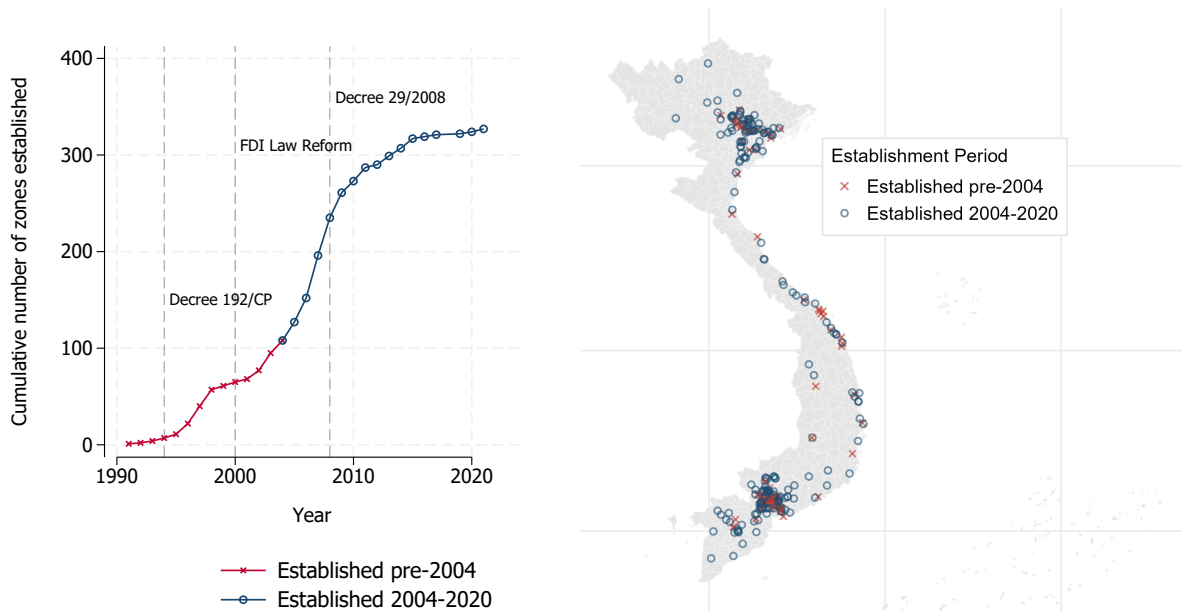
The right panel of Figure 1 illustrates a spatial shift in industrial development. Zones established before 2004 were largely concentrated around the country's major economic centers (e.g., Hanoi, Ho Chi Minh City, Can Tho, and Da Nang), reflecting early priorities to build on existing urban and coastal infrastructure. In contrast, zones established from 2004 onward are more geographically dispersed, extending into less developed regions such as the northwest, central coast, Central Highlands, and deeper areas of the Mekong River Delta. This spatial diffusion reflects a deliberate policy shift toward more balanced regional development and provides the identifying

---

<sup>3</sup>Firms in zones located in areas classified as having "difficult" or "especially difficult" socioeconomic conditions qualify for more generous incentives: a preferential CIT rate of 10% for up to 15 years, with four years of full exemption and nine subsequent years at 50% reduction.

<sup>4</sup>Law on Import and Export Duties (No. 45/2005/QH11, replaced by No. 107/2016/QH13) and Decree 134/2016/ND-CP. Export-oriented firms are additionally exempt from duties on imported raw materials used in the production of exported goods.

Figure 1: Growth and Spatial Distribution of Industrial Zones in Vietnam Over Time

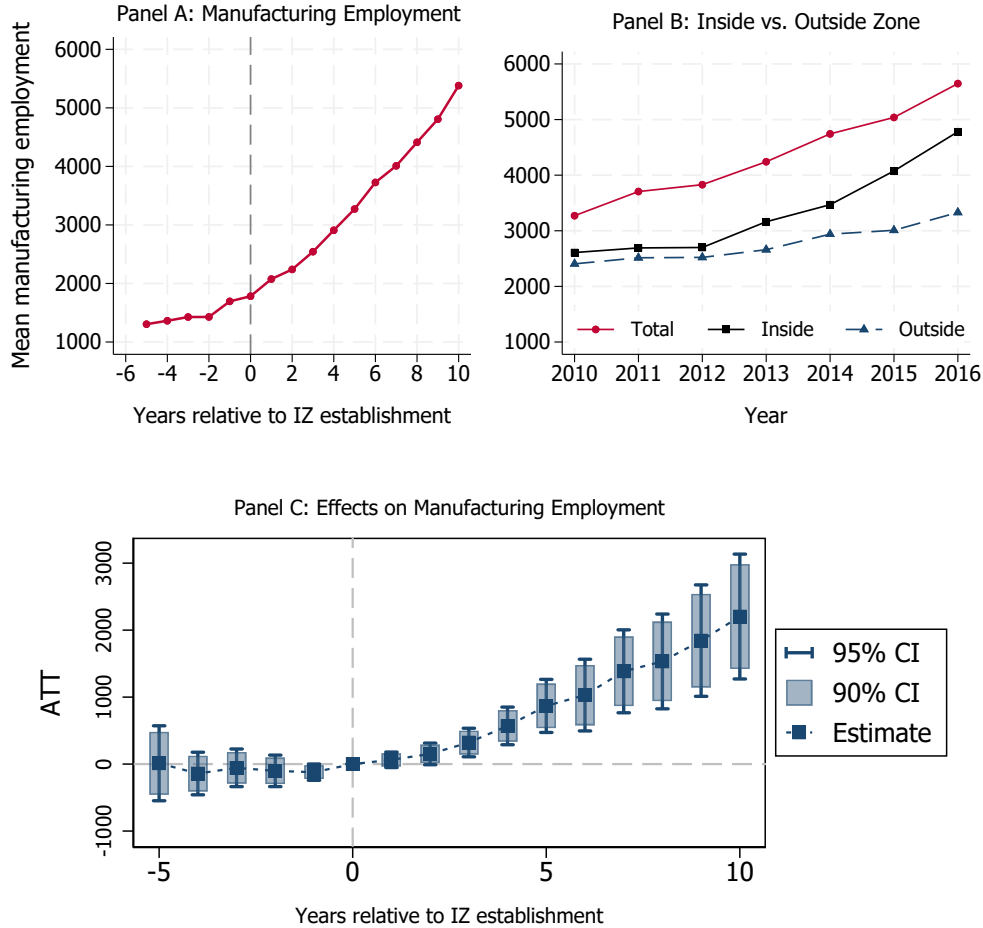


*Notes:* This figure presents two panels. The left panel shows the cumulative number of established and operational industrial zones in Vietnam by year, distinguishing zones established before 2004 (red x markers) from those established during the study period of 2004 to 2020 (navy circles). The right panel displays the spatial distribution of industrial zones using the same markers and colors. Data on individual zones are sourced from the Ministry of Planning and Investment.

variation exploited in the empirical analysis.

Industrial zones have expanded local manufacturing activity. Figure 2 provides evidence using district-level data from the Vietnam Enterprise Survey (VES), a comprehensive annual census of formal enterprises conducted by the General Statistics Office of Vietnam. Panel A plots average manufacturing employment per district in event time, aligning each district within 15 km of a zone to its own establishment year. Manufacturing employment is relatively stable in the years before zone establishment and accelerates sharply afterward, increasing from around 1,300 to more than 5,000 workers per district within a decade. Panel B decomposes manufacturing employment within zone-hosting districts into firms located inside and outside the zone boundary, using data from VES 2010–2016 when firm-level zone classification is available. Both components grow over this period: inside-zone employment nearly doubles from around 2,600 to 4,800 workers, while outside-zone employment increases by roughly 40%, suggesting that zones generate new manufacturing activity rather than simply displacing it from surrounding areas. Panel C presents the corresponding event-study estimates of Panel A, obtained using the de Chaisemartin and d’Haultfoeuille (2024) estimator with district and year fixed effects and baseline district char-

Figure 2: Industrial Zones and Manufacturing Expansion



Notes: Panel A plots the mean manufacturing employment across treated districts (within 15 km of an industrial zone) aligned to their respective establishment years (event time = 0). Panel B restricts to treated districts for which firm-level zone classification is available and plots mean manufacturing employment per district separately for firms located inside and outside the zone boundary. Panel C presents the corresponding event-study estimates of Panel A, using the de Chaisemartin and d’Haultfoeuille (2024) estimator with district and region-by-year fixed effects, and baseline district characteristics interacted with linear time trends, standard errors clustered at the district level. Source: Vietnam Enterprise Survey, 2001 to 2018 (Panels A and C) and 2010 to 2016 (Panel B).

acteristics interacted with linear time trends. Pre-treatment coefficients are close to zero, supporting the parallel trends assumption, and the post-treatment path confirms that the acceleration in manufacturing employment coincides with zone establishment.

Table 1 characterizes the types of jobs created by industrial zones, comparing firms inside and outside zones within the same district using data from VES 2016. Three features stand out. First, industrial zones are overwhelmingly manufacturing-oriented: manufacturing accounts for approximately 65% of firms inside zones and 83% of zone employment, compared with roughly one-sixth of firms outside zones. Zones also attract a markedly higher share of foreign-invested

firms, nearly half compared with only about 2% outside zones, reflecting the program’s role in channeling foreign direct investment into export-oriented production.

Table 1: Firm and Employment Characteristics, Outside and Inside Zones

	Outside	Inside	Difference	
	Zones	Zones	(1) – (2)	
	Mean	Mean	Coef.	p-value
	(1)	(2)	(3)	(4)
Share of Foreign-Invested Firms	0.015	0.447	0.432	0.000
Share of Firms with 100+ Employees	0.028	0.330	0.302	0.000
Share of Firms in Manufacturing	0.133	0.649	0.515	0.000
Share of Employment in Manufacturing	0.367	0.830	0.463	0.000
Share of Workers Without High School Diploma*	0.493	0.491	-0.002	0.919
Share of Manufacturing Workers Without High School Diploma*	0.563	0.514	-0.049	0.023
Share of Manufacturing Workers With a College Degree**	0.140	0.207	0.067	0.001
Average Wage Bill of Manufacturing Workers (Million 2010 VND)	4.343	4.898	0.556	0.000

*Notes:* Columns (1) and (2) report mean values for firms located outside and inside zones, respectively. Column (3) reports the coefficient on an inside-zone dummy controlling for district fixed effects; standard errors are clustered at the district level. Column (4) reports the corresponding p-value. The first three rows are unweighted firm-level shares. The share of employment in manufacturing and the worker education shares are weighted by firm employment. The wage bill is defined as total payments to workers divided by the number of workers, weighted by firm employment, and restricted to manufacturing firms. \*Workers classified as untrained, having less than three months of training, or possessing elementary-level or intermediate-level vocational training are considered not to have completed high school education. \*\*Workers with a college/university/Master’s/PhD degree are considered completing college education. *Source:* Calculations from Vietnam Enterprise Survey 2016.

Second, the overall share of workers without a high school diploma is nearly identical inside and outside zones (approximately 49%), but differences emerge within manufacturing. The share of manufacturing workers without a high school diploma is about 5 percentage points lower inside zones (51%) than outside (56%), while the college-educated share remains small in both locations (21% versus 14%). The relevant skill threshold for zone employment is therefore upper secondary completion, not post-secondary education. This is consistent with the occupational structure of manufacturing: a World Bank employer survey finds that approximately one-third of manufacturing occupation titles require less than high school education, involving tasks such as simple machine operation, elementary assembly, or freight handling, while 43% require at least upper secondary education but not a university degree, including more technical roles such as technicians, assemblers, and quality control workers (Granata et al., 2023). Upper secondary completion therefore opens access to a substantially broader range of zone jobs, motivating the education-based sample splits used in the empirical analysis.

Third, wages are higher inside zones. The average monthly wage bill per manufacturing worker is approximately 4.9 million VND inside zones, compared with 4.3 million VND outside zones, a difference of about 14%. Both are substantially higher than agricultural wages of approximately 2.8 million VND. This wage premium suggests that industrial zones can influence local economic conditions not only through employment creation but also through higher earnings among manufacturing workers, with potential demand spillovers to surrounding areas. The extent to which these spillovers reach households with limited access to formal zone employment is the central empirical question of this paper.

## 2.2 Labor Market Context

During the same period as rapid industrial zone expansion, Vietnam’s labor market experienced substantial changes in workforce composition and the sectoral distribution of employment (Table 2). Youth labor force participation declined over time, falling from 71% in 2004 to 66% in 2020, while participation among prime-age adults remained high and stable.

Table 2: Labor Market Statistics, 2004–2020

	Labor Participation		Sectoral Labor Share Without High School Diploma			Sectoral Labor Share With High School Diploma		
	Age 19-24	Age 25-64	Agriculture	Informal Non-Ag	Formal Non-Ag	Agriculture	Informal Non-Ag	Formal Non-Ag
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2004	0.71	0.92	0.60	0.31	0.08	0.23	0.27	0.50
2012	0.64	0.91	0.53	0.36	0.11	0.18	0.26	0.56
2020	0.66	0.90	0.34	0.46	0.20	0.09	0.30	0.61

*Notes:* This table reports the share of individuals in each age group who worked in the last 12 months (Columns 1–2), as well as the sectoral distribution of working individuals by educational attainment—without a high school diploma (Columns 3–5) and with a high school diploma (Columns 6–8), based on the primary (most time-consuming) occupation. “Informal” is defined as employment in small-scale household businesses or self-employment. Sampling weights are applied throughout. *Source:* Calculations using VHLSS 2004, 2012, 2020.

Changes in the sectoral composition of employment differ sharply by educational attainment. For workers without a high school diploma, the dominant shift was out of agriculture (from 60% to 34%) and into informal non-agricultural activities (from 31 to 46%), with a more modest increase in formal employment (from 8 to 20%). For workers with at least a high school diploma, the decline in agriculture (from 23 to 9%) was absorbed almost entirely by formal non-agricultural employment (from 50 to 61%), while informal non-agricultural work remained stable. These pat-

terns suggest pronounced educational stratification in Vietnam’s structural transformation: less-educated workers transition primarily into informal activities, while more-educated workers move into formal employment.

This stratification becomes particularly relevant when considering potential spillovers from industrial zones. The expanding informal non-farm sector, which consists largely of elementary occupations in trade, food services, and accommodation, provides employment opportunities for less-educated workers who are excluded from formal-sector jobs. This sector is also highly dynamic: annual entry and exit rates of 14 to 18% reflect low barriers to entry and a high degree of fluidity across activities, particularly relative to self-employment in agriculture (McCaig & Pavcnik, 2021). Its accessibility to less-educated workers, low entry costs, and operational flexibility make it a plausible channel through which the benefits of industrial zones may extend beyond direct formal employment.

### **3 CONCEPTUAL FRAMEWORK AND TESTABLE PREDICTIONS**

This paper examines how the establishment of industrial zones reshapes local labor allocation and household welfare. A central feature of the setting is that formal employment expands but remains unevenly accessible across education groups. To structure the empirical analysis and guide interpretation, I develop a stylized local-economy framework with three production activities (agriculture, formal non-agriculture, and informal non-agriculture) and two household types that differ in their access to formal non-farm employment. The framework highlights two key features: (i) industrial zones increase formal production and earnings; and (ii) local spillovers operate through non-tradable goods and services that are produced and consumed locally.

The model’s main implication is that industrial zones affect less-educated households not only through their limited direct access to formal jobs, but also through an indirect local-demand channel. Higher formal earnings, accruing disproportionately to more-educated households, raise demand for local services, increase the equilibrium returns to informal activities, and induce labor reallocation away from agriculture. Importantly, this adjustment need not take the form of a clean transition from one sector to another. When formal employment is limited and informal opportunities are expanding, less-educated households may spread labor across multiple activities, tem-

porarily diversifying their labor portfolios across agriculture, formal, and informal work. This within-household diversification is a policy-relevant margin of adjustment: if it is large enough, individual-level employment patterns that appear to favor educated workers may mask more inclusive welfare gains at the household level. These mechanisms generate distinctive patterns of sectoral adjustment, household diversification, and welfare that guide the empirical tests. Appendix A formalizes this framework, derives the comparative statics, and provides conditions under which these predictions hold.

### 3.1 Environment

**Endowment.** Consider a local economy with a continuum of households of two skill types: high-skill ( $H$ ), and low-skill ( $L$ ). Each household allocates one unit of labor across three activities: agriculture ( $A$ ), formal non-agriculture ( $F$ ), and informal non-agriculture ( $I$ ) so that:

$$l_A^s + l_F^s + l_I^s = 1, \quad s \in \{H, L\} \quad (1)$$

**Production and Segmented Access to Formal Employment.** Agriculture and the formal non-farm sector produce tradable output sold at exogenous prices, while informal non-farm activities produce non-tradable goods and services consumed locally at an equilibrium local price  $p_I$ . Household enterprises in agriculture and informal non-farm activities feature diminishing returns to labor, reflecting small-scale production and limited complementary inputs.

Industrial zone exposure is summarized by a scalar  $\theta$  that increases the scale and productivity of the formal tradable sector. In the simplest formulation, higher  $\theta$  raises formal wages for both types, with a skill premium:

$$w_F^H(\theta) > w_F^L(\theta), \quad \frac{dw_F^s(\theta)}{d\theta} > 0 \quad (2)$$

Here  $\theta$  can be interpreted as the arrival or expansion of zone firms, infrastructure, and demand for formal non-farm labor.

A key feature of the framework is that formal non-farm employment is not freely accessible for

low-skill households. Instead, their participation is limited by an access ceiling:

$$l_F^L \leq \bar{l}(\theta, \phi) \quad (3)$$

where  $\phi$  is the local skill intensity (i.g., the share of high-skill households in the local population). The ceiling increases with  $\theta$ , which reflects the idea that industrial zones create additional formal non-farm jobs, but decreases with  $\phi$ , capturing stronger competition and screening in more skill-intensive places. A convenient reduced-form representation is:

$$\bar{l}(\theta, \phi) = \frac{\bar{l}_0 \theta}{1 + \phi} \quad (4)$$

which can be micro-founded by a simple rationing or screening mechanism: zone expansion increases the number of formal non-farm positions and higher-skill applicants have a higher effective probability of being selected. In the analysis, I focus on a range where this constraint binds for low-skill households, consistent with persistent barriers to formal work even after zones are established.

High-skill households are assumed to have substantially better access to formal jobs and, once zones are sufficiently important, allocate nearly all labor to formal employment. This captures the idea that the primary adjustment margin for high-skill households is increased formal work and formal income, while the main question of interest is how low-skill households adjust when formal access expands only partially.

**Preference.** Households allocate spending between a tradable composite and non-tradable goods and services. With homothetic preferences (e.g., Cobb-Douglas), non-tradable spending is a constant share of income, so aggregate demand for non-tradables is increasing in total local income and decreasing in the local price  $p_I$ :

$$D_I(\theta, p_I) = \frac{\mu Y^{total}(\theta, p_I)}{p_I}, \quad \mu \in (0, 1) \quad (5)$$

**Equilibrium.** Non-tradable supply  $Y_I(\theta, p_I)$  is produced locally through informal activities, and in this stylized setting is primarily supplied by low-skill households. The local non-tradables

market clears when:

$$D_I(\theta, p_I) = Y_I(\theta, p_I) \quad (6)$$

This equilibrium condition is the central link between industrial zones and informal non-farm sector returns. When  $\theta$  rises, two forces increase excess demand for non-tradables at the initial price. First, higher formal wages raise local income (especially among high-skill households with better access to formal work), increasing demand for local services. Second, low-skill households reallocate some labor into formal work as the access ceiling relaxes, which can reduce the labor available for informal production. These demand and supply forces move in the same direction, implying that the equilibrium non-tradable price  $p_I$  rises with  $\theta$ . Intuitively, this is a local-economy analog of Balassa–Samuelson and local multiplier mechanisms: formal-sector expansion raises local demand for services and bids up their price when supply is local (Balassa, 1964; Moretti, 2010; Samuelson, 1964).

**Implications for Low-Skill Reallocation and Diversification.** The rise in  $p_I$  increases the return to informal activities relative to agriculture, inducing low-skill households to shift non-formal labor away from agriculture and toward informal non-farm activities. Low-skill adjustment therefore operates along two margins:

- (i) Direct formal margin: low-skill formal work increases as access relaxes, although it remains capped by  $\bar{l}(\theta, \phi)$ .
- (ii) Indirect non-formal margin: within non-formal labor, low-skill households substitute from agriculture into informal non-farm activities as informal returns increase with  $p_I$ .

This two-margin structure implies a non-monotonic pattern for household labor diversification. When low-skill households start agriculture-heavy and formal access is limited, small increases in formal work combined with a shift into informal services move households toward a more balanced allocation across sectors, raising diversification. As exposure increases further, informal services can absorb most non-formal labor; at that point, additional shifts toward informal work may concentrate the non-formal portfolio and diversification can flatten or decline. This mechanism motivates the empirical focus not only on sectoral shares, but also on household labor portfolios and diversification measures.

**Welfare Implications.** A natural welfare measure is real income, which depends on both nominal earnings and the local non-tradable price. The model highlights that rising  $p_I$  is not simply a cost-of-living increase for low-skill households. Low-skill households are the primary producers of non-tradables in the local economy, while high-skill households are primarily demanders. As a result, low-skill households are net sellers of non-tradables in equilibrium. An increase in the non-tradable price therefore raises low-skill households' informal earnings and can generate welfare gains even as the local price level increases. This implies that industrial zones can improve low-skill welfare through both (i) direct formal earnings where access expands and (ii) indirect gains through the local-demand channel that increases informal returns.

### 3.2 Testable Predictions

The framework yields several testable predictions that guide the empirical analysis:

- (i) Formal non-farm sector effects: Industrial zone establishment increases formal non-farm employment and earnings, with larger gains for high-skill households due to higher wages and better access. Low-skill formal participation increases but remains constrained.
- (ii) Non-tradable spillovers: zone exposure raises local returns to informal non-farm activities.
- (iii) Reallocation within low-skill labor: low-skill households shift non-formal labor away from agriculture and toward informal non-agriculture as non-tradable returns rise.
- (iv) Diversification dynamics: low-skill diversification increases in agriculture-heavy settings at low-to-moderate exposure, but may flatten or decline as informal activities become dominant, implying a non-monotonic relationship between exposure and labor portfolio concentration.

In the model, diversification arises because a household spreads its single unit of labor across three activities. In practice, this can occur through two channels: different household members working in different sectors (across-member diversification), or individual members combining jobs in multiple sectors (within-member diversification). Both channels reflect the same underlying mechanism, namely that when formal access is limited and informal returns are rising, households find it optimal to maintain a presence in multiple activities

rather than concentrating in one. In the empirical analysis, I measure diversification along both margins, as well as through a continuous index based on the distribution of total household labor hours across sectors.

- (v) Heterogeneity by skill intensity: the model suggests stronger non-tradable spillovers and larger low-skill portfolio adjustments in more skill-intensive districts, where formal income gains (and thus local demand shifts) are larger and where low-skill households remain more reliant on non-formal work.

## 4 DATA AND EMPIRICAL STRATEGY

Guided by the conceptual framework, this section describes the data and empirical strategy used to estimate the impact of industrial zones and discusses potential threats to identification.

### 4.1 Data and Construction of Variables

**Data.** The first primary source of data is a comprehensive industrial zone database from the Ministry of Planning and Investment, containing detailed records for each zone: name, administrative location (ward, district, province), establishment date, operational status as of June 2023, and key performance indicators such as the number of domestic and foreign secondary projects. As the database did not include geographic coordinates, I manually georeferenced each zone using Google Maps.

The identification strategy relies on changes at the district level. Districts are Vietnam’s second-level administrative units under provinces. To map zones to districts, I used district-level administrative boundary shapefiles from the Humanitarian Data Exchange (COD dataset), reflecting 2019–2020 boundaries.<sup>5</sup> Because administrative boundaries changed over the study period through district splits and mergers, I harmonized the spatial units by aggregating newly created subdivisions back to their original “mother districts.” This process yields a consistent panel of approximately 660 districts, compared to over 700 in the unadjusted 2020 boundaries. I then spatially joined the georeferenced industrial zones to these harmonized district units.

I use ten rounds of the Vietnam Household Living Standards Survey (VHLSS), conducted bi-

---

<sup>5</sup><https://data.humdata.org/dataset/cod-ab-vnm>

ennially by Vietnam’s General Statistics Office from 2002 to 2020. The survey employs a stratified multi-stage cluster sampling design, first selecting communes or wards (the third-level administrative units below provinces and districts) stratified by urban or rural status and geographic region, then randomly selecting households within these primary sampling units. Each wave surveys more than 45,000 households across over 3,000 communes, designed to be representative at national, regional, urban/rural, and provincial levels. While the large sample provides substantial coverage for district-level analysis, I address potential limitations in district-level representativeness through robustness checks including province-level estimation where representativeness is assured. All analyses incorporate appropriate sampling weights.

**Construction of Variables.** The main outcomes are individual measures of employment, as well as household measures of labor diversification and income.

*Employment Classification.* For all individuals aged 10 and above, the VHLSS Employment Module collects information on participation in income-generating activities over the past 12 months, including household farm and non-farm work as well as wage employment. From 2004 onward, respondents report the industry, employer type, and hours worked for their two most time-consuming jobs, whereas the 2002 round records only the primary job. Because measures of labor diversification and sector-specific hours require information on secondary employment, the main analysis uses the 2004–2020 survey rounds.<sup>6</sup> I restrict the sample to persons aged 25 to 64 to focus on the working-age population, excluding most students and retirees.

Following McCaig and Pavcnik (2018), I classify employment as informal if it consists of self-employment or work in household businesses, including cooperatives and collectives for consistency over time.<sup>7</sup> Based on their primary job, I classify each worker into agriculture, formal non-agriculture, or informal non-agriculture. I also calculate total hours in each sector using both primary and secondary jobs.

*Household Labor Diversification.* Using individual-level employment data, I construct three measures of household labor diversification to capture different dimensions of how labor is distributed

---

<sup>6</sup>The exception is the measure of skill intensity, which is constructed using 2002 data.

<sup>7</sup>This enterprise-based definition (formal registration status) correlates highly (Pearson coefficient  $\approx 0.9$ ) with worker-based definitions using social insurance and contracts (General Statistics Office of Vietnam and International Labor Organization, 2018). Cooperatives are included as informal because they have historically operated outside the formal enterprise system, with collectives participation being mostly voluntary and informal (Nguyen et al., 2016; Raymond, 2008). In addition, cooperatives are legally distinct from formal enterprises under the 2012 Cooperative Law.

both across and within households. The first measure, which directly corresponds to the model, is the *sectoral diversification index*, a continuous variable based on the Herfindahl index that quantifies the overall distribution of household labor. It is calculated as:

$$\text{Index} = 1 - \sum_j \left( \frac{\text{hours}_j}{\sum_j \text{hours}_j} \right)^2$$

where  $j$  indexes sectors (agriculture  $A$ , formal non-agriculture  $F$ , and informal non-agriculture  $I$ ), and  $\text{hours}_j$  represents total household hours in sector  $j$ . Values closer to 1 indicate greater labor diversification across sectors; values near 0 imply concentration in a single sector.

The second measure, *across-member diversification*, is a binary indicator equal to 1 if household members are employed in different sectors (e.g., one member works in agriculture while another works in formal or informal non-agriculture) and 0 if all members work exclusively within the same sector. This captures sectoral variety across household members.

The third measure, *within-member diversification*, is a binary variable equal to 1 if any household member is employed in more than one sector (e.g., holding a primary job in formal non-agriculture and a secondary job in another sector), and 0 if all members have jobs in only one sector. This captures multi-sector employment within individuals.

Together, these measures provide a comprehensive view of household labor diversification, capturing both the spread of employment across household members and sectors as well as multi-sector engagement at the individual level.

*Household Labor and Business Income.* I measure household income by aggregating earnings across all household members. For wage workers, the survey directly reports labor compensation, including base wages, bonuses, incentives, and allowances. For self-employed workers, I proxy labor compensation using the VHLSS Business Module, which records revenues and costs for all household farm and non-farm enterprises. I compute sector-specific household profits as revenues minus costs, including payments to hired workers. All income measures are deflated to 2010 Vietnamese Dong using the national CPI to ensure comparability across sectors and over time.

*Complementary Welfare Measures.* As additional evidence on whether income gains translate into tangible improvements in living standards, I examine household consumption, human capital spending, and child schooling outcomes. The VHLSS collects detailed consumption data from

a representative stratified sub-sample of approximately 9,000 households per wave. Food consumption combines regular food and drink items consumed in the past 30 days (annualized) and holiday food expenditure over the past 12 months, including both purchased and self-supplied items. Non-food consumption comprises recurring daily non-food items (reported over the past 30 days and annualized) and annual non-food items such as clothing, household goods, and communications, excluding food, housing, education, and health expenditure. I also examine combined spending on education (tuition, textbooks, uniforms, and tutoring) and health (in-patient and out-patient treatment). All consumption and spending measures are expressed in per capita terms and deflated to 2010 Vietnamese Dong. For children aged 10–18, I measure school enrollment using the Education Module, which records whether each child attended school in the past 12 months. These outcomes are available for smaller sub-samples, which limits statistical power, but they provide a useful check on whether the income gains documented in the main analysis extend to household consumption and human capital investment.

I merge individual- and household-level VHLSS data with the geospatial database of industrial zones at the district level. While the VHLSS does not report exact household locations, the district identifiers allow me to determine which households are exposed to nearby industrial zones based on the harmonized district boundaries. Specifically, a district is classified as treated if its boundary falls within 15 km of an industrial zone, capturing the commuting range over which zones are likely to affect local labor markets. I discuss this treatment definition in Section 4.3.

## 4.2 Summary Statistics

Table 3 presents baseline characteristics of sample districts in 2004, measured from the VHLSS. The sample includes districts within 15 km of a zone established from 2004 onward (173 districts) and districts that were never within 15 km of any zone during the study period (220 districts). The remaining 268 districts (those within 15 km of a zone established before 2004) are excluded, as they lack a pre-treatment period for identification. Columns (1)–(3) compare never-treated districts with those within 10 km of a zone established during the study period. Ever-treated districts are more urban (16% versus 13%), have lower ethnic minority shares (16% versus 38%), and are less reliant on agricultural labor (62% versus 73%). These differences are consistent with non-random zone placement in areas with greater industrial potential.

Table 3: Characteristics of Sample Districts in 2004

	Never-treated Districts	Districts with Zones since 2004	Difference (2) – (1)	Districts with Zones 2004–2008	Districts with Zones 2009–2013	Districts with Zones 2014–2020	Difference (5) – (4)	Difference (6) – (4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Demographics</i>								
Share of urban population	0.126 [0.178]	0.158 [0.217]	-0.032 (0.022)	0.148 [0.204]	0.179 [0.248]	0.211 [0.266]	0.032 (0.049)	0.064 (0.078)
Share of ethnic minority population	0.381 [0.403]	0.159 [0.260]	0.222 (0.038)	0.130 [0.229]	0.286 [0.345]	0.128 [0.222]	0.156 (0.061)	-0.002 (0.055)
<i>Panel B: Adult Labor Outcomes</i>								
Share of agricultural labor	0.732 [0.174]	0.623 [0.200]	0.109 (0.021)	0.635 [0.183]	0.618 [0.248]	0.511 [0.222]	-0.016 (0.046)	-0.124 (0.067)
Share of formal non-agricultural labor	0.095 [0.074]	0.110 [0.085]	-0.015 (0.008)	0.102 [0.078]	0.126 [0.105]	0.154 [0.092]	0.024 (0.019)	0.052 (0.027)
Share of informal non-agricultural labor	0.173 [0.139]	0.267 [0.154]	-0.094 (0.017)	0.263 [0.145]	0.256 [0.188]	0.334 [0.143]	-0.007 (0.036)	0.071 (0.044)
Income from wages and HH business profits	2.319 [0.738]	2.496 [0.624]	-0.177 (0.082)	2.463 [0.601]	2.633 [0.697]	2.509 [0.676]	0.170 (0.138)	0.046 (0.211)
<i>Panel C: Household Labor and Agriculture</i>								
HH labor diversification: across members	0.476 [0.156]	0.538 [0.158]	-0.062 (0.018)	0.545 [0.162]	0.491 [0.152]	0.580 [0.102]	-0.054 (0.030)	0.035 (0.037)
Total sown area for crop production (ha)	1.106 [0.844]	1.035 [0.763]	0.071 (0.094)	0.976 [0.668]	1.457 [1.050]	0.604 [0.370]	0.481 (0.203)	-0.373 (0.127)
Share of rice area	0.535 [0.260]	0.622 [0.248]	-0.087 (0.031)	0.630 [0.236]	0.600 [0.288]	0.595 [0.284]	-0.030 (0.057)	-0.034 (0.098)
Share of industrial crop area	0.121 [0.178]	0.101 [0.181]	0.020 (0.021)	0.116 [0.192]	0.064 [0.154]	0.029 [0.059]	-0.053 (0.035)	-0.088 (0.023)
<i>Panel D: Child Schooling</i>								
Share of 10-18 attending school	0.790 [0.122]	0.778 [0.105]	0.012 (0.014)	0.783 [0.100]	0.748 [0.118]	0.803 [0.103]	-0.035 (0.025)	0.020 (0.031)

*Notes:* This table summarizes 2004 characteristics across district types. Columns (1)–(3) compare never-treated versus ever-treated districts: Column (1) includes districts that have not received nor in proximity to any zone during the study period; Column (2) includes those within 15 km of a zone established in 2004 or later; Column (3) reports mean differences. Columns (4)–(7) compare treatment timing within the study period: Column (4) includes districts within 15 km of a zone established 2004–2008; Column (5) includes those within 15 km of a zone established 2009–2013; Column (6) includes those within 15 km of a zone established 2014–2020. Columns (7)–(8) report mean differences. Standard deviations in brackets; robust standard errors in parentheses. Sampling weights applied throughout.

*Source:* Data from VHLSS 2004.

Columns (4)–(8) compare districts by the timing of zone establishment. In contrast to the level differences between ever-treated and never-treated districts, baseline characteristics are largely balanced across adoption cohorts. Formal and informal non-agricultural employment, household income, school enrollment, and crop composition are all statistically indistinguishable across the three waves. Some variation exists in agricultural labor shares and total sown area, but these differences do not follow a systematic pattern across cohorts. Nonetheless, to the extent that baseline differences across cohorts generate heterogeneous treatment effects, the main analysis employs heterogeneity-robust estimators, which accommodates such heterogeneity rather than imposing the homogeneity assumption implicit in standard TWFE.

### 4.3 Empirical Strategy

As Table 3 shows, districts that received industrial zones differ systematically from those that did not, making simple comparisons uninformative about causal effects. To address this non-random placement, I exploit the staggered rollout of zones across districts over time in a difference-in-differences framework.

**Staggered DiD Design.** I estimate an event-study specification as follows:

$$y_{idt} = \alpha_d + \alpha_t + \sum_{k \neq 0}^K \beta_k \cdot \mathbb{I}(t = \text{Establishment}_d + k) + \mathbf{X}'_d \phi \cdot t + \varepsilon_{idt} \quad (7)$$

where  $y_{idt}$  is the outcome of individual or household  $i$  in district  $d$  at time  $t$ . District fixed effects  $\alpha_d$  absorb district-specific time-invariant characteristics such as geography, soil quality, distance to major cities, and historical patterns of economic development that may be correlated with both zone placement and labor market outcomes. Year fixed effects  $\alpha_t$  absorb nationwide time-varying shocks common to all districts, such as changes in macroeconomic conditions, national trade policy, or commodity prices. The term  $\mathbf{X}_d \cdot t$  interacts baseline district characteristics—the share of urban population, ethnic minority population, agricultural labor force, and rice sown area, measured in 2002 or, for districts not surveyed in that round, in 2004—with a linear time trend, allowing districts with different initial compositions to follow differential trajectories that would otherwise confound the treatment effects. Since all zone establishments in the study sample occur from 2004 onward, these characteristics are pre-determined with respect to treatment.

Identification therefore comes from comparing outcome changes over time between treated and control districts, after removing common nationwide shocks, time-invariant district characteristics, and differential trends associated with observable baseline conditions. The coefficients  $\beta_k$  capture the dynamic treatment effects  $k$  periods relative to zone establishment. The omitted period is  $k = 0$ , normalizing effects relative to the year of treatment. Standard errors are clustered at the district level to account for spatial and temporal correlation.

**Treatment Definition.** I define a district as treated if any part of its territory falls within 15 kilometers of an industrial zone’s geographic centroid. This distance-based measure captures the localized nature of zone impacts while maintaining a clear, replicable treatment assignment rule (Duranton & Venables, 2018; Gallé et al., 2024). The 15-kilometer threshold is motivated by two considerations. First, existing evidence suggests that economic spillovers from place-based policies such as special zones typically concentrate within 10-15 kilometers (e.g., Abagna et al., 2025; Gallé et al., 2024). Second, my own analysis confirms this pattern: examining outcomes at distance bins from zone centers shows that employment and welfare effects attenuate sharply beyond 15 kilometers (see Appendix Section C).

Treatment begins when a zone is “established,” that is, when its non-infrastructure projects receive official approval from authorities. While zones may not be fully operational at this point, establishment marks the transition from planning into tangible development activity, triggering expectations among investors, workers, and local governments, as well as prompting other economic responses such as employment adjustments, land use changes, and early migration. Given that delays between establishment and operation are common due to administrative or logistical bottlenecks, using the establishment date as the onset of treatment provides a more realistic measure of when local economic dynamics begin to respond to zone development. To ensure consistency and avoid bias from zones that were established but never materialized, the analysis is limited to zones that became fully operational by the end of the study period.<sup>8</sup>

---

<sup>8</sup>According to the zone database, only one zone had been terminated as of June 2023 due to delays in infrastructure development. Appendix Figure B1 illustrates the staggered expansion of treatment over time for a province. Appendix Table B1 provides a summary of the staggered establishment of industrial zones across districts throughout the study period. Zone establishment was most active during the early years, with approximately 10% of districts establishing zones during 2004–2008. Zone creation then declined substantially, with only 0.4% during 2009–2020. The treatment exposure patterns follow a similar trajectory. Notably, more than 35% of districts were never exposed to any industrial zone throughout the entire study period, providing substantial comparison group for the analysis.

**Identifying Assumptions.** For the coefficients  $\beta_k$  to be interpreted as causal effects of industrial zones, the following assumptions must hold. First, the parallel assumption requires that if no industrial zone had been established, labor market and household welfare outcomes would have evolved similarly on average across all district groups—those that established a zone at different times and those that never did. Second, the no anticipation assumption requires that a district’s outcomes in a given period when untreated are unaffected by the timing of future treatment; in other words, individuals and households in these districts do not adjust their behavior in anticipation of future zone establishment before it formally occurs.

While the parallel trends assumption is fundamentally untestable, the event-study specification allows me to assess its plausibility by examining pre-treatment trends. The coefficients  $\beta_k$  where  $k < 0$  capture differences in trends between treated and control districts before zone establishment (relative to  $k = 0$ ). A lack of (joint) statistical significance of these coefficients would indicate no systematic pre-treatment trend differences, suggesting that the parallel trends assumption held before treatment and may plausibly continue to hold. This test also supports the no anticipation assumption by showing no anticipatory behavioral adjustments. To assess the sensitivity of the results to potential violations of this assumption, I follow Rambachan and Roth (2023) and construct confidence sets that remain valid under the assumption that post-treatment violations of parallel trends are no larger than the maximum violation observed in the pre-treatment period. This approach provides an informative bound on the estimated effects even when pre-trend tests lack statistical power.

A related concern is that agricultural shocks, such as commodity price fluctuations or extreme weather events, could differentially affect districts depending on their crop composition, generating divergent trends unrelated to zone establishment. While year fixed effects absorb common shocks affecting all districts in a specific year, differential exposure remains a concern if districts vary in crop composition. However, treated districts across adoption cohorts exhibit similar crop composition at baseline (Table 3, Panel C), making differential exposure to crop-specific shocks unlikely. In addition, the baseline specification controls for the share of rice sown area interacted with a linear time trend, allowing districts with different agricultural compositions to follow differential trajectories.

In addition, the Stable Unit Treatment Value Assumption (SUTVA) requires that one district’s

treatment status does not affect outcomes in other districts. This could be violated if zones generate spillovers to neighboring districts or if they induce migration that affects control areas. I address this through spatial ring analysis showing that effects attenuate sharply beyond 15 km (Appendix Section C), a donut exclusion test removing potentially contaminated districts (Table B7, Row 9), and sample restrictions to long-term registered residents (Table B7, Row 7). Results from all three exercises support the baseline findings.

**Further Considerations.** Three additional issues merit discussion. First, recent advances in the DiD literature have demonstrated that conventional TWFE estimation can produce biased estimates when treatment effects vary across cohorts and over time, as it implicitly uses already-treated units as controls for later-treated units, and in some cases, applies potentially negative weights to some comparisons (e.g., Borusyak et al., 2024; de Chaisemartin & d’Haultfoeuille, 2020; Goodman-Bacon, 2021). Treatment effect heterogeneity is plausible in my setting: zones established at different times serve different economic purposes (United Nations Industrial Development Organization, 2019; World Bank Group, 2019), and some baseline differences exist across adoption cohorts (Table 3). Applying the diagnostic decomposition of Goodman-Bacon (2021) confirms that only 75% of the TWFE weight is assigned to clean comparisons using never-treated districts as controls, whereas the remaining 25% relies on comparisons using already-treated or earlier-treated districts as controls.

To address these issues, I estimate event-study effects using the heterogeneity-robust difference-in-differences framework of de Chaisemartin and d’Haultfoeuille (2024), rather than estimating equation (7) with standard OLS. For each treated district, this approach compares the outcome evolution from the period before treatment to  $\ell$  periods after, between that district and districts whose treatment has not yet changed, ensuring that already-treated units are never used as controls. It accommodates treatment effect heterogeneity across cohorts and over time, and uses placebo estimators for pre-treatment periods to assess the parallel trends and no-anticipation assumptions. I report two pre-treatment periods and up to six post-treatment periods.<sup>9</sup> An addi-

---

<sup>9</sup>Extending to three pre-treatment periods restricts the estimation sample to 21 switching districts from cohorts treated after 2010, which account for fewer than 10% of all treated districts. These late cohorts are not representative of the broader set of cohorts that drive the treatment effect estimates, making the longer-horizon placebo tests unreliable as a pre-trend diagnostic. By comparison, the first and second placebo periods draw on 98 and 49 switching districts, respectively.

tional advantage of this framework is that it accommodates non-binary treatments, which I exploit in province-level robustness checks where treatment is measured as a continuous variable reflecting the intensity of zone exposure. I also report results using the alternative heterogeneity-robust estimator of Callaway and Sant'Anna (2021), as well as event-study TWFE estimates. All yield qualitatively similar results.

Second, because the VHLSS is designed to be representative at the provincial rather than district level, district-level estimates may be sensitive to sampling variation. I address this through several checks: re-estimating at the province level using the cumulative number of zones as a continuous treatment variable, re-estimating on household panels exploiting the rotating panel structure of the VHLSS, and performing leave-one-district-out and Monte Carlo permutation exercises. All yield conclusions consistent with the baseline.

Finally, treatment and outcome timing do not perfectly align: zones can be established in any year, while household surveys occur biennially. In the baseline, a district's treatment status in a given survey year reflects all zones established up to and including that year. Zones established between survey waves are therefore first captured in the subsequent wave, while zones established in a survey year are treated as contemporaneous. This means that the baseline may attenuate estimated effects for zones established in survey years, as the household data may partially predate zone activity. Restricting the sample to households surveyed at least one full survey cycle after zone establishment does not change the findings (Section 6).

## 5 EMPIRICAL FINDINGS

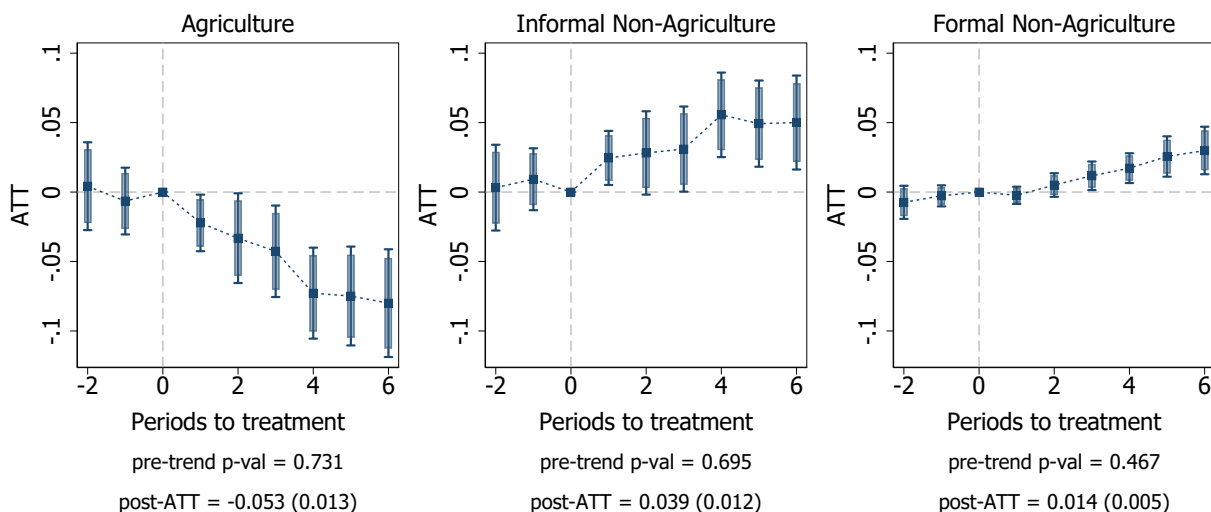
This section presents the empirical estimates of the impact of industrial zones on individual and household-level labor responses and welfare outcomes. As shown in Section 2, zone establishment is associated with a sharp increase in local manufacturing firms and employment, with an average of 10 additional firms and 1,000 new jobs within ten years of establishment (Figure 2 and Appendix Figure B2), providing the labor demand shift that drives the adjustments documented below.

## 5.1 Labor Market Adjustments: Participation and Sectoral Allocation

Before examining sectoral reallocation, I confirm that industrial zones do not significantly affect overall labor force participation, defined as whether an individual engaged in any income-generating activity in the past 12 months, among prime-age and older individuals (25–64), with post-treatment effects near zero (Appendix Table B2). While overall labor force participation remains stable, the sectoral composition of employment shifts significantly in response to industrial zone establishment.

Figure 3: Industrial Zones and Primary Sector Employment, Ages 25–64

Sample includes individuals without a high school diploma



*Notes:* This figure shows the effects of industrial zone exposure on sectoral employment shares among all individuals aged 25–64 without a high school diploma, using data from VHLSS 2004–2020. The outcome is the sector of an individual’s primary occupation (the most time-consuming job in the past 12 months). Square markers indicate the point estimates of the coefficients. Darker vertical lines with caps show 95% confidence intervals, and lighter bars represent 90% confidence intervals. *pre-trend* p-val is the p-value from the joint test that pre-treatment effects are zero. *post-ATT* represents the average treatment effect on the treated across post-treatment periods, with standard errors clustered at the district level in parentheses. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

Figure 3 presents the dynamic effects of industrial zone exposure on the sectoral distribution of workers’ primary employment, with a focus on adults without a high school diploma. Pre-treatment coefficients are small and statistically indistinguishable from zero across sectors, supporting the parallel trends assumption. Exposure to industrial zones is associated with a gradual decline in the share whose primary job is in agriculture. The effect becomes more pronounced several periods after zone establishment, reaching a reduction of approximately 7–8 percentage

points by four to five periods (i.e., 8 to 10 years) post-treatment. Over the same horizon, the decline in agricultural employment is accompanied by increases in both informal and formal non-agricultural employment, with the expansion much larger in the informal sector on average.

Table 4 examines the robustness of these patterns using two additional employment measures and compares responses across education groups. Panel A classifies workers by whether they hold a job in a given sector as either their primary or secondary activity, while Panel B restricts to workers whose sole employment is in a single sector. Comparing these measures helps distinguish between complete exit from a sector, the addition of new activities alongside existing work, and the dropping of secondary activities. This is an important distinction in settings where multi-activity livelihoods are common (e.g., Barrett et al., 2001; Fields, 2011; Fox & Gaal, 2008).

Table 4: Industrial Zones and Sector Employment: Heterogeneity by Education Level

Sector	Without High School Diploma			With High School Diploma		
	Agriculture	Informal Non-Ag	Formal Non-Ag	Agriculture	Informal Non-Ag	Formal Non-Ag
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Primary or Secondary Job</i>						
post-ATT	-0.042 (0.013)	0.045 (0.012)	0.014 (0.005)	-0.038 (0.023)	-0.010 (0.020)	0.053 (0.021)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.78	0.38	0.07	0.41	0.33	0.52
p-value pre-trend	0.966	0.472	0.223	0.790	0.815	0.257
<i>Panel B: Only Job (Single Sector)</i>						
post-ATT	-0.056 (0.013)	0.029 (0.012)	0.009 (0.003)	-0.029 (0.017)	-0.001 (0.016)	0.025 (0.019)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.56	0.19	0.03	0.19	0.18	0.37
p-value pre-trend	0.381	0.922	0.876	0.174	0.555	0.800
N(Individual-Period)	226471	226471	226471	53136	53136	53136
N(District Switcher-Period)	838	838	838	826	826	826

Notes: This table shows the effects of industrial zone exposure on sectoral employment shares among working individuals aged 25–64, using data from VHLSS 2004–2020. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Sampling weights are applied throughout.

Among less-educated workers, the sectoral patterns documented in Figure 3 are robust across both measures. Agricultural employment declines by 4.2 percentage points when secondary jobs are included (Panel A, Column 1;  $p < 0.01$ ) and by 5.6 percentage points for exclusive agricultural

employment (Panel B, Column 1;  $p < 0.01$ ), consistent with the 5.3 percentage point decline in primary agricultural employment shown in Figure 3. Formal non-agricultural employment rises by 0.9–1.4 percentage points across panels ( $p < 0.01$ ), representing sizable gains relative to low baseline levels of 3–7%. Informal non-agricultural employment increases by a larger 3.0–4.6 percentage points ( $p < 0.01$ ), corresponding to 10–15% increases relative to baseline shares of 19–38%. Notably, the presence of significant effects in both the “primary or secondary job” and “only job” measures indicates that informal employment expansion reflects entry as a primary source of livelihood rather than merely a supplemental activity alongside agricultural work. The consistency of the agricultural decline across all three measures (primary job, any job, and sole job) suggests that the reduction reflects genuine exits from the sector rather than a reclassification of agriculture from primary to secondary activity.

For individuals with at least a high school diploma, the adjustment pattern differs in both magnitude and destination. Prior to zone exposure, this group was already substantially less engaged in agriculture—25% in primary jobs, 41% in primary or secondary jobs, and 19% working exclusively in agriculture—compared with 66%, 78%, and 56%, respectively, among less-educated workers. Industrial zone exposure reduces agricultural employment by 3.8 percentage points when secondary activities are included (Panel A, Column 4) and by 2.8 percentage points for exclusive agricultural employment (Panel B, Column 4). These effects, however, are not statistically significant at conventional levels.

The key distinction from less-educated workers lies in the destination sector. Among more-educated individuals, the entire reallocation flows into formal non-agricultural employment, which increases by 2.5–5.3 percentage points across panels ( $p < 0.05$ ), corresponding to roughly 6–9% gains relative to baseline. Informal non-agricultural employment shows no significant change (Column 5). This contrasts sharply with the pattern for less-educated workers, where the informal sector accounts for the majority of the transition out of agriculture. The divergence is consistent with the model’s formal access constraint: both education groups gain formal non-agricultural employment, but the likelihood of accessing formal jobs is roughly two to three times higher for workers with a high school diploma, a differential reflected in the relative magnitudes across Columns 3 and 6. With limited formal access, less-educated workers reallocate their remaining labor toward informal non-agricultural activities.

At face value, these patterns could suggest divergent welfare implications: more-educated workers gain access to higher-quality formal employment, while less-educated workers disproportionately enter informal activities with potentially lower returns. However, such an interpretation overlooks an important margin of adjustment in low- and middle-income settings, i.e., household-level labor diversification. Households may respond by reallocating labor across members or by combining multiple activities across sectors. If less-educated households use diversification strategies to offset limited access to formal employment, the distributional consequences of industrial zone development may be more inclusive than individual-level employment patterns alone would imply.

## 5.2 Household-Level Labor Diversification

I begin by examining the sectoral diversification index (one minus the Herfindahl index of household labor hours across sectors), which corresponds directly to the diversification measure in the conceptual framework (Proposition A3). Figure 4 plots event-study estimates for households without a high school graduate. Pre-treatment coefficients are close to zero and jointly insignificant ( $p = 0.84$ ), supporting the parallel trends assumption. After zone establishment, diversification rises gradually, becoming clearly positive within a few post-treatment periods and peaking around periods 4–5 before leveling off. This pattern is consistent with the model’s non-monotonic prediction: when less-educated households are initially agriculture-heavy, expanding formal access and rising informal returns increase the spread of labor across sectors, but as informal activities absorb an increasing share of non-formal labor, diversification gains diminish.

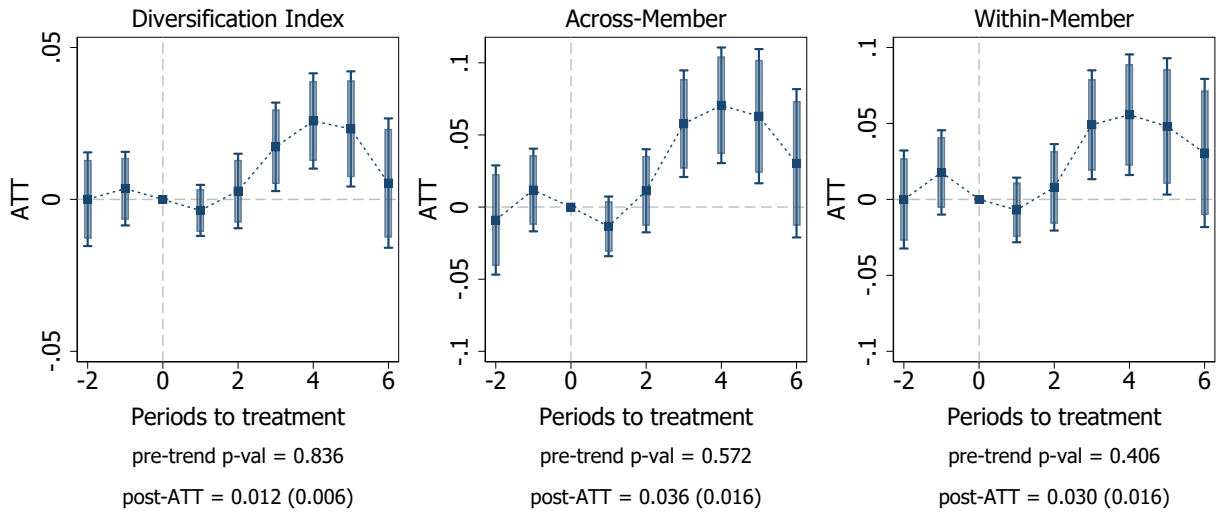
To provide more insights into the diversification dynamics, Appendix Figure B3 left panel plots event-study estimates for the distance of the household labor portfolio from an equal split of hours shares across agriculture, formal non-agriculture, and informal non-agriculture:

$$\text{distance-to-balance} = \sum_{j \in \{A, I, F\}} (l_j - 1/3)^2$$

Because distance-to-balance is minimized at  $l_A = l_I = l_F = 1/3$ , a decline in this measure provides a direct summary of movement toward the diversification-maximizing portfolio. Following industrial zone exposure, distance-to-balance declines steadily and reaches its lowest point

Figure 4: Industrial Zones and Household Labor Diversification Index

Sample includes households without high school graduates



Notes: This figure shows the effects of industrial zone exposure on different measures of household labor diversification, using data from VHLSS 2004–2020. The square symbol represents the point estimates of the coefficients. Darker vertical lines with caps indicate 95% confidence intervals, while the lighter bars represent 90% confidence intervals. *pre-trend* p-val is the p-value from the joint test that pre-treatment effects are zero. *post-ATT* represents the average treatment effect on the treated across post-treatment periods, with standard errors in parentheses. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

around periods 4–5, implying that households move away from an initially concentrated portfolio toward a more even distribution of labor across activities. At longer horizons the effect attenuates, distance-to-balance rebounds toward zero by period 6, mirroring the hump-shaped pattern in the diversification measures. Together with the household-level sectoral hours responses (Appendix Figure B4), these results are consistent with the model mechanism in which zone expansion initially spreads labor away from agriculture and toward non-agriculture, but later adjustments yield smaller gains in portfolio balance.<sup>10</sup>

Table 5 extends the analysis to other measures of labor diversification, reporting period-specific estimates alongside the average post-treatment estimate. For less-educated households, all three diversification measures are near zero at period 1 and peak at period 4: the sectoral index rises to 0.026 ( $p < 0.01$ ), across-member diversification by 7 percentage points ( $p < 0.01$ ), and within-member diversification by 6 percentage points ( $p < 0.01$ ). By period 6, the effects attenuate: the

<sup>10</sup>Note that the number of working household members remains relatively stable (Appendix Table B3), indicating that diversification reflects reallocation of existing workers across sectors rather than an increase in household labor supply.

sectoral index falls to 0.005, and the binary measures decline by roughly half, which are consistent with the model's prediction that diversification gains diminish as informal activities absorb an increasing share of non-formal labor. The response operates through both members specializing in different sectors and individuals combining activities across sectors.

Households with at least one high school graduate also show increased across-member diversification (Column 5: 0.05 at period 4,  $p < 0.05$ ), but no within-member response (Column 6: -0.005 at period 6, insignificant). Two features distinguish this group's response from that of less-educated households. First, whereas less-educated households diversify along both the across- and within-member margins, within-member diversification (individual members working multiple sectors) is absent among households with a high school graduate. This is the margin most directly linked to the model's compensating mechanism, in which skill-constrained workers combine activities across sectors when formal access is limited. Second, unlike less-educated households, where diversification attenuates at longer horizons consistent with the model's non-monotonic prediction, the across-member response for households with a high school graduate remains stable over time (0.046 at period 4, 0.033 at period 6), suggesting a persistent compositional shift rather than a transitional adjustment strategy.

To interpret these differential responses, it is worth noting that the empirical classification of households with some high school graduates includes those with at least one high school graduate, who typically contain both educated and less-educated members, the modal composition in rural Vietnam. Columns 7–9 sharpen this interpretation by restricting to households where all working-age members hold a high school diploma. This group shows no diversification response on any margin: point estimates are small and insignificant at period 4 (Column 7: -0.005; Column 8: 0.002; Column 9: 0.014) and, if anything, turn negative at longer horizons. The contrast between Columns 4–6 and 7–9 confirms that the across-member diversification observed among mixed-skill households reflects differential sorting of heterogeneous members into different sectors, with the educated member taking a formal job while less-educated members remain in agriculture or informal work, rather than the compensating diversification strategy the model predicts for skill-constrained households. When all members have direct access to formal employment, no diversification response emerges.

Table 5: Industrial Zones and Household Labor Diversification  
Heterogeneity by Household Education Level

	Without High School Graduate			Some High School Graduates			All High School Graduates		
	(1) Sectoral Index	(2) Across Member	(3) Within Member	(4) Sectoral Index	(5) Across Member	(6) Within Member	(7) Sectoral Index	(8) Across Member	(9) Within Member
Period1-ATT	-0.004 ( 0.004)	-0.013 ( 0.011)	-0.007 ( 0.011)	-0.003 ( 0.007)	-0.009 ( 0.014)	-0.003 ( 0.015)	-0.011 ( 0.011)	-0.043 ( 0.026)	-0.038 ( 0.022)
Period4-ATT	0.026 ( 0.008)	0.071 ( 0.020)	0.056 ( 0.020)	0.024 ( 0.010)	0.051 ( 0.022)	0.031 ( 0.023)	-0.005 ( 0.013)	0.002 ( 0.032)	0.014 ( 0.032)
Period6-ATT	0.005 ( 0.011)	0.030 ( 0.026)	0.031 ( 0.025)	0.023 ( 0.011)	0.046 ( 0.024)	-0.005 ( 0.029)	-0.033 ( 0.015)	-0.069 ( 0.038)	-0.034 ( 0.042)
post-ATT	0.012 ( 0.006)	0.036 ( 0.016)	0.030 ( 0.016)	0.017 ( 0.008)	0.033 ( 0.016)	0.017 ( 0.019)	-0.012 ( 0.011)	-0.026 ( 0.026)	-0.004 ( 0.027)
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean Outcome	0.18	0.48	0.35	0.26	0.63	0.40	0.15	0.39	0.27
p-value pre-trend	0.836	0.572	0.406	0.616	0.456	0.180	0.934	0.767	0.251
N(Household-Period)	92518	92518	92518	53893	53893	53893	19890	19890	19890
N(District Switcher-Period)	839	839	839	832	832	832	792	792	792

*Notes:* This table shows the effects of industrial zone exposure on household-level labor diversification strategies, across households without any high school graduates, with some high school graduates, and with all working-age members having high school diploma. In Columns (1), (4) and (7), the dependent variable calculated as 1 minus the Herfindahl-Hirschman Index (HHI) using the share of household labor hours in each sector (higher values indicate greater diversification). In Columns (2), (5), and (8), the dependent variable is whether the household has different members working in different sectors. In Columns (3), (6) and (9), the dependent variable is whether the household has at least one member working in different sectors. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d'Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Sampling weights are applied throughout.

The broader pattern of household diversification documented here is consistent with a long literature on multi-activity labor patterns in developing economies. Such patterns are often interpreted as income smoothing and risk management in settings with incomplete insurance and credit markets (Barrett et al., 2001; Dercon, 2002; Morduch, 1995; Rosenzweig & Stark, 1989; Udry, 1996). An alternative interpretation, emphasized in the labor and informality literature, is that diversification can arise from limited access to stable formal employment, leading households to combine activities across sectors when high-quality jobs are limited (Harris & Todaro, 1970; Maloney, 2004; Ulyseas, 2020). The within-member results provide evidence favoring the latter mechanism: it is specifically the margin through which constrained individuals combine activities that responds among less-educated households but not among those with formal access through an educated member. The gradient across all three household types, from strong diversification with no educated members, to across-member sorting only with some educated members, to no response when all members are educated, traces out the model's formal access constraint directly.

Cross-district variation in the diversification response provides further evidence on which interpretation is more consistent with the data. Under a risk-diversification view, one might expect the response to be at least as strong in areas with lower skill intensity, where informal income tends to be more volatile and insurance markets thinner. Under the formal access view formalized in Section 3, the response would instead be stronger in areas with higher skill intensity, where formal income gains are larger and may generate greater local demand for informal services (Corollary A1). The pattern in the data appears more consistent with the latter: diversification effects among less-educated households are substantially larger in districts above the 75th percentile of baseline skill share compared with those below, with magnitudes two to four times as large depending on the measure (Appendix Table B4 and Figure B3 middle-right panels). That said, because baseline skill composition is correlated with other local characteristics such as urbanization and initial formalization, I consider this pattern as suggestive rather than as a sharp test of the mechanism.

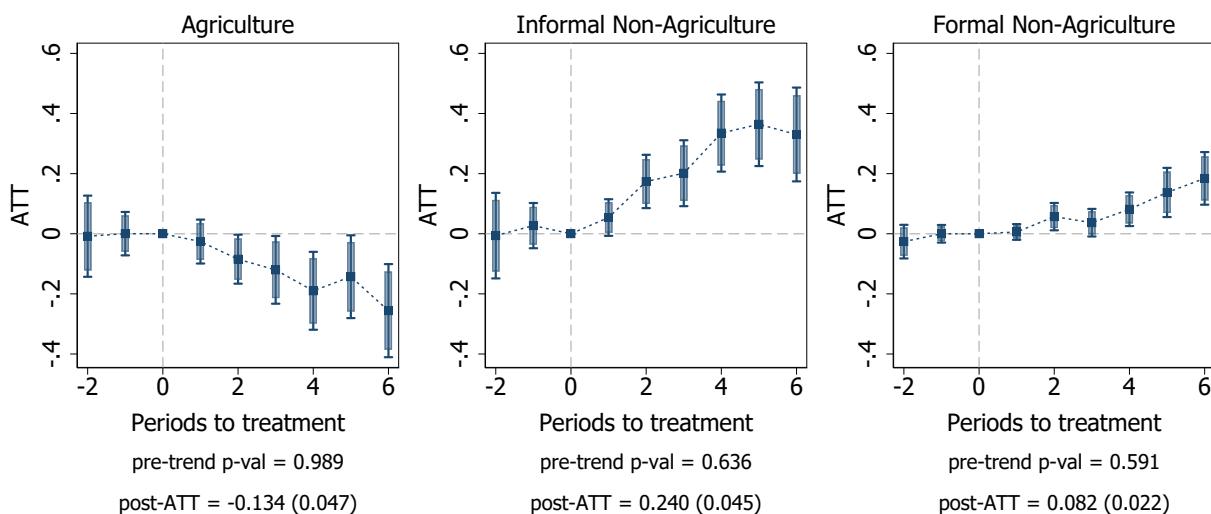
### **5.3 Economic Gains: Income, Consumption, and Human Capital Investment**

The sectoral reallocation and diversification patterns documented above raise a central question for evaluating industrial zones in high-informality settings: do these labor adjustments translate into broad-based income gains, or do benefits concentrate among households able to access formal

jobs? Guided by the conceptual framework, I next examine household income by source (farming, informal non-farm, and formal non-farm activities) to trace how zone-induced labor reallocation maps onto household earnings.

Figure 5: Industrial Zones and Household Labor Income

Sample includes households without high school graduates



*Notes:* This figure shows the effects of industrial zone exposure on household income (monthly) from labor and business profits, by source, using data from VHLSS 2004–2020. Household income from each sector is the sum of individual wage earnings aggregated to the household level and household-level business profits (net revenues minus operating costs). Sector-level income is winsorized at the 2.5th and 97.5th percentiles within each survey year, education group, and treatment status, conditional on positive participation; households not engaged in a given sector are assigned zero income for that source. All monetary values are expressed in millions of 2010 VND. The square symbols represent point estimates of the coefficients. Darker vertical lines with caps indicate 95% confidence intervals, while the lighter bars represent 90% confidence intervals. pre-trend p-val is the p-value from the joint test that pre-treatment effects are zero. post-ATT represents the average treatment effect on the treated across post-treatment periods, with standard errors clustered at the district level in parentheses. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

Figure 5 presents event-study estimates for each income source among less-educated households. Pre-treatment coefficients are close to zero and jointly insignificant for all three outcomes (p-values of 0.98, 0.66, and 0.49), supporting the parallel trends assumption. Three patterns emerge after the treatment. First, farming income declines by 0.134 million VND ( $p < 0.01$ ), consistent with the reallocation of labor out of agriculture. Second, informal non-farm income rises by 0.240 million VND ( $p < 0.01$ ), more than offsetting the farming income decline. Third, formal non-farm income increases more modestly at 0.082 million VND ( $p < 0.01$ ), consistent with the limited formal sector access available to this group. Informal non-farm income thus accounts for 75% of total non-farm income gains, with formal income contributing the remaining 25%.

To examine whether the informal income gains are consistent with the local demand channel emphasized in the model, Table 6 disaggregates informal business activity among less-educated households into non-tradable and tradable sectors. If zone establishment raises local demand for non-tradable goods and services, business formation and performance should increase in non-tradable activities but not in tradable ones. The results are consistent with the local demand channel. The probability of operating a non-tradable business increases by 3.3 percentage points (Column 1,  $p < 0.01$ ), representing a 15% increase relative to the pre-treatment mean. Non-tradable business revenue and profits also increase, respectively, by 1.144 (Column 2,  $p > 0.10$ ) and 0.753 million VND (Column 3,  $p < 0.10$ ). In contrast, tradable business activity shows no significant response: the probability of operation is unchanged (Column 4), and revenue and profit effects are smaller and not statistically significant at conventional levels (Columns 5–6). This pattern confirms that informal income gains are driven by activities serving local markets rather than by a general improvement in the business environment near zones.

Table 6: Industrial Zones and Informal Household Non-Farm Business Performance

	Heterogeneity by Activities					
	Non-Tradable Sector			Tradable Sector		
	(1) Business (0/1)	(2) Revenue	(3) Profits	(4) Business (0/1)	(5) Revenue	(6) Profits
post-ATT	0.033 (0.011)	1.144 (0.731)	0.753 (0.421)	-0.002 (0.006)	0.358 (0.437)	0.068 (0.169)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.24	8.14	4.87	0.07	2.10	0.97
p-value pre-trend	0.460	0.654	0.544	0.553	0.634	0.537
N(Household-Period)	96363	96363	96363	96363	96363	96363
N(District Switcher-Period)	839	839	839	839	839	839

*Notes:* This table shows the effects of industrial zone exposure on informal business activity among less-educated households, separately for non-tradable and tradable industries. Non-tradable industries include construction, retail, wholesale, food services, accommodation, transport, personal services, repair services, healthcare, and education. Tradable industries include manufacturing, mining, and other activities producing goods for external markets. Columns (1) and (4) report effects on the probability of operating at least one business in the respective category. Columns (2)–(3) and (5)–(6) report effects on household-level profits and revenue, which include zeros for households not operating a business in that category. Profits and revenue are winsorized at the 5th and 95th percentiles within each survey year, education group, and treatment status, conditional on positive values. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Sampling weights are applied throughout.

Furthermore, the model predicts that the local demand channel is stronger in districts with higher skill intensity, where more-educated workers generate larger formal income gains and thus greater demand for non-tradable services (Corollary A1). Appendix Table B5 provides suggestive evidence consistent with this prediction: non-tradable business formation among less-educated households is approximately three times larger in high-skill districts (8.4 percentage points) than in low-skill districts (2.6 percentage points), and non-tradable revenue and profit effects are also larger in high-skill districts, though less precisely estimated.

Table 7: Industrial Zones and Household Labor Income

Sector	Heterogeneity by Education Level					
	Without High School Graduates			With High School Graduates		
	Agriculture	Informal Non-Ag	Formal Non-Ag	Agriculture	Informal Non-Ag	Formal Non-Ag
	(1)	(2)	(3)	(4)	(5)	(6)
post-ATT	-0.134 (0.047)	0.240 (0.045)	0.082 (0.022)	-0.095 (0.075)	0.030 (0.092)	0.413 (0.096)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	1.38	0.79	0.16	1.30	1.30	1.38
p-value pre-trend	0.989	0.636	0.591	0.492	0.512	0.030
N(Household-Period)	96363	96363	96363	54657	54657	54657
N(District Switcher-Period)	839	839	839	832	832	832
Share of non-ag income gain (%)		74.53	25.47		6.77	93.23

*Notes:* This figure shows the effects of industrial zone exposure on household income (monthly) from labor and business profits, by source, using data from VHLSS 2004–2020. Household income from each sector is the sum of individual wage earnings aggregated to the household level and household-level business profits (net revenues minus operating costs). Sector-level income is winsorized at the 5th and 95th percentiles within each survey year, education group, and treatment status, conditional on positive participation; households not engaged in a given sector are assigned zero income for that source. All monetary values are expressed in millions of 2010 VND. pre-trend p-val is the p-value from the joint test that pre-treatment effects are zero. post-ATT represents the average treatment effect on the treated across post-treatment periods, with standard errors clustered at the district level in parentheses. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

Table 7 extends the income analysis to households with at least one high school graduate. For this group, the composition of income gains reverses: formal non-farm income accounts for 93% of non-farm gains (0.41 million VND,  $p < 0.01$ ), while informal non-farm income is small and not statistically significant (0.03 million VND).<sup>11</sup> This divergence in income composition mirrors the

<sup>11</sup>The pre-trend test for formal non-farm income among more-educated households is significant ( $p = 0.030$ ) and thus this result should be interpreted with caution, though the direction of the effect is consistent with expectations and with the labor allocation results, which show clean pre-trends for this group.

labor allocation results, where more-educated households shift primarily into formal employment while less-educated households diversify into informal activities.

Taken together, these results suggest that assessments of place-based policies based solely on formal labor market outcomes would capture most of the income gains for more-educated households but only a quarter for less-educated households. In settings where informal employment is prevalent, such an approach risks understating the welfare impact for the group with the least access to formal jobs.

As complementary measures of household welfare, I examine three outcomes for less-educated households: food and non-food consumption expenditure per capita, combined health and education spending per capita, and school enrollment among children aged 10–18. These outcomes test whether the income gains documented above translate into tangible improvements in living standards and human capital investment. Point estimates are positive across all measures: food consumption increases by 2.2% (not statistically significant), non-food consumption increases by 10% ( $p < 0.01$ ), health and education spending rises by 15% ( $p < 0.01$ ) relative to pre-treatment mean values, and school enrollment increases by roughly 2 percentage points ( $p < 0.05$ ). These results, reported in Appendix Table B6, are consistent with the local demand channel generating welfare improvements that extend beyond income to household consumption and investment in children’s schooling.

## 6 ROBUSTNESS CHECKS

The credibility of the staggered difference-in-differences design rests on the parallel trends assumption. While the event-study figures show no evidence of differential pre-trends, a concern remains that trends could diverge non-linearly after treatment. I assess this using the framework of Rambachan and Roth (2023), which constructs robust confidence intervals under progressively larger departures from a linear extrapolation of the pre-treatment trend. The parameter  $M$  governs the degree of permitted non-linearity:  $M = 0$  imposes that any violation of parallel trends follows a linear path extrapolated from the pre-treatment period, and larger values of  $M$  allow for increasingly non-linear deviations. As shown in Appendix Figure B5, the core findings on sectoral reallocation, household labor diversification, informal income gains, and business formation

are robust both under the linear extrapolation ( $M = 0$ ) and across the range of larger  $M$  values considered.<sup>12</sup>

The main findings are robust to a range of specification and inference checks (Table B7). Extending the comparison group to include not-yet-treated districts alongside never-treated districts yields virtually identical estimates (Row 2), alleviating concerns that never-treated districts may differ systematically from treated districts. Clustering standard errors at the province level leaves all conclusions unchanged (Row 3). Re-estimating with the doubly-robust estimator of Callaway and Sant’Anna (2021) yields qualitatively consistent results across all outcomes (Row 4). Row 5 reports event-study estimates from a conventional TWFE specification. The estimates are qualitatively consistent with the baseline across all outcomes but attenuated in magnitude, as expected given that approximately 25% of the TWFE weight is assigned to comparisons using already-treated or earlier-treated districts as controls (Section 4). Adding individual-level (or household head-level) demographic controls for age, gender, and ethnic minority status does not affect the results (Row 6), and restricting the sample to long-term residents with permanent household registration in the same commune produces near-identical estimates (Row 7), addressing concerns that observed effects reflect migrant composition rather than local responses. To address the potential timing mismatch arising from the biennial structure of the VHLSS, I restrict the sample to households surveyed at least one full survey cycle after zone establishment, and the findings remain unchanged (Row 8).

Three additional checks address concerns specific to the spatial and sampling structure of the data. To assess whether spillovers from zone activity contaminate the control group, I exclude districts located 15–25 km from the nearest industrial zone. The resulting estimates are larger in magnitude across all outcomes (Row 9), consistent with positive spillovers attenuating the baseline contrast rather than driving it, and suggesting that the baseline estimates are if anything conservative. To assess whether the results are robust to tracking the same households over time rather than relying on repeated cross-sections, I re-estimate the baseline specification on a household panel constructed by pooling four overlapping three-wave rotating panels (2004–2008, 2010–2014, 2012–2016, and 2014–2018), in which the same households are surveyed across three consecutive

---

<sup>12</sup>There is no single threshold for  $M$  that determines robustness and the interpretation depends on the economic plausibility of the implied non-linearities (Rambachan & Roth, 2023).

biennial waves. The estimates remain qualitatively consistent, though with wider confidence intervals reflecting the smaller effective sample (Row 10). Furthermore, to address the concern that the VHLSS is designed to be representative at the province rather than the district level, I aggregate the analysis to the province level using the cumulative number of zones as a continuous treatment variable following de Chaisemartin and d’Haultfoeuille (2024) and de Chaisemartin et al. (2024). The results remain qualitatively consistent with the baseline, though smaller in magnitude as expected given that province-level estimates capture the average effect across all households including those far from any zone (Row 11).

Finally, I conduct two diagnostic exercises to assess the sensitivity and validity of the baseline estimates. Leave-one-district-out estimation, in which each ever-treated district is sequentially excluded and the baseline specification re-estimated, confirms that no single district exerts undue influence on the results. The leave-one-out coefficients are tightly distributed around the baseline values across all six outcomes (Appendix Figure B6). A Monte Carlo permutation test, in which the treatment status time series is randomly reassigned across districts 500 times, confirms that the baseline estimates fall far outside the resulting distribution of placebo coefficients, with observed Type I error rates close to the nominal 5% level and randomization-based p-values below 0.01 for all outcomes (Appendix Figure B7).

## 7 DISCUSSION

### 7.1 Migration and the Local Demand Channel

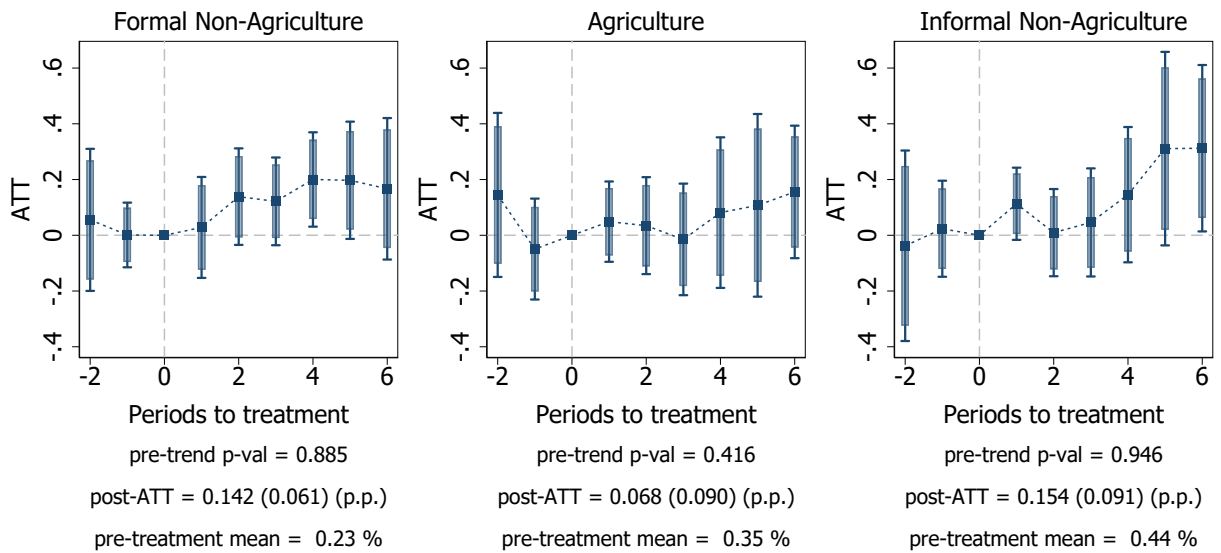
Industrial zones may affect the local economy not only through direct labor demand but also through the migration of workers into treated districts. If zones attract formal workers from outside the area, these workers’ local spending could itself be a source of the increased demand for non-tradable services that benefits incumbent less-educated households. Understanding the role of migration is therefore important for interpreting how the demand channel operates.

There are reasons to expect limited migration in this context. The districts in the sample, those receiving zones from 2004 onward, are predominantly located in less urbanized, agricultural areas (Table 3), in contrast to the earlier wave of zones concentrated in major economic hubs such as Ho Chi Minh City, Hanoi, and Binh Duong, where industrial expansion attracted large inflows of

migrant workers. The baseline share of residents without long-term household registration (in the same commune) is just above 1%, suggesting that these areas were not significant migration destinations prior to zone establishment.

Notwithstanding these priors, there are some evidence of a selective migration response. I construct indicators for whether an individual of working age (25–64) is both non-registered in the commune of residence and employed in a given sector, and estimate the baseline specification using each indicator as an outcome. As shown in Figure 6, industrial zone establishment increases the share of non-local formal workers by 0.14 percentage points ( $p < 0.05$ ), a 62% increase relative to the pre-treatment mean of 0.23%. The effect builds gradually, consistent with progressive hiring as zone firms expand. There is no corresponding inflow of agricultural workers, confirming that the migration response is specific to the labor demand created by zone firms. The share of non-local informal workers shows a positive but imprecisely estimated trajectory in later periods, consistent with downstream attraction as the local service economy matures.

Figure 6: Industrial Zones and Share of Migrant Workers, Ages 25–64



Notes: This figure shows the effects of industrial zone exposure on the share of working-age individuals (25–64) who are employed in a given sector (for their primary job) and not registered in the commune of residence, using data from VHLSS 2004–2020. The sample includes all education levels. Square markers indicate the point estimates of the coefficients. Darker vertical lines with caps show 95% confidence intervals, and lighter bars represent 90% confidence intervals. *pre-trend* p-val is the p-value from the joint test that pre-treatment effects are zero. *post-ATT* represents the average treatment effect on the treated across post-treatment periods, with standard errors clustered at the district level in parentheses. *pre-treatment mean* reports the mean of the outcome variable in treated districts prior to zone establishment. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

These patterns suggest that in-migration of formal workers is part of the mechanism through which zones generate local demand spillovers. Importantly, the welfare gains documented for less-educated households are not driven by compositional change: restricting the sample to long-term registered residents yields near-identical estimates (Appendix Table B7, Row 5), confirming that the benefits accrue to the incumbent population through the local spending of both resident and migrant workers (Table 6).

## 7.2 Implications for Place-Based Policy Evaluation

Despite the rapid global expansion of special economic zones, particularly across South and Southeast Asia and Sub-Saharan Africa (UNCTAD, 2019), the evidence on whether these programs improve welfare in surrounding communities remains mixed (Abagna et al., 2025; Busso et al., 2013; Gallé et al., 2024; Kline & Moretti, 2014; Y. Lu et al., 2019; Rothenberg et al., 2025). The findings in this paper suggest that two conditions help explain this heterogeneity.

The first condition is that zones must successfully generate formal economic activity. As documented in Figure 2, industrial zone establishment in Vietnam led to a sharp expansion in manufacturing firms and employment within treated districts. This formal-sector expansion is the prerequisite for all downstream effects: without it, there is no wage income to generate local demand, and the informal spillover channel cannot operate. In Indonesia, Rothenberg et al. (2025) find that industrial zones had no significant effect on local economic outcomes, a result consistent with zones that failed to attract sufficient firm entry. The modest effects documented for US enterprise zones (Busso et al., 2013) may similarly reflect limited firm creation relative to the pre-existing local economy. Even within programs that do raise employment, the response may be uneven: Criscuolo et al. (2019) find that UK investment subsidies increased manufacturing jobs only among small firms, with large firms accepting subsidies without expanding activity. By contrast, the large effects found in China (Y. Lu et al., 2019; Wang, 2013) and the structural transformation documented in India (Gallé et al., 2024) are consistent with settings where zones generated substantial formal activity. In Vietnam, Tafese et al. (2025) find that SEZ expansion shifts employment from agriculture into formal manufacturing in foreign firms, with earnings spillovers to workers in domestic firms. The present paper documents a distinct channel: rather than higher earnings within existing informal activities, industrial zones increase the formation and performance of

non-tradable businesses and expand informal income for less-educated households through local demand for goods and services.

The second condition is that the local economy must have the capacity to transmit formal-sector gains through informal channels. In Vietnam, the informal non-farm sector consists largely of elementary occupations in trade, food services, and accommodation, with annual entry and exit rates of 14 to 18% reflecting low barriers to entry and high operational fluidity (McCaig & Pavcnik, 2021). These features make the sector particularly responsive to local demand shocks. The results in this paper confirm that this sector is the primary transmission channel: informal income gains account for approximately 75% of the total non-farm income effect for less-educated households, concentrated in non-tradable services (Table 6). In economies with tight labor markets and high formality, this absorptive capacity is likely absent, and formal-sector demand shocks are more likely to raise local prices or dissipate into the broader market rather than generating localized multiplier effects.

The structural features that enable the informal demand channel in Vietnam, including agricultural labor surpluses, low-barrier informal service sectors, and skill-based segmentation of formal employment, are shared by many of the economies where zone programs are expanding most rapidly. Recent evidence from 10 African countries documents household wealth gains within 10 km of SEZs, accompanied by occupational shifts away from agriculture (Abagna et al., 2025). These patterns are consistent with the local demand mechanism documented here, though the specific income channels remain to be identified.

## 8 CONCLUSION

This paper argues that standard approaches to evaluating place-based industrial policies are systematically incomplete in economies where a large share of the workforce operates informally. Existing evaluations typically measure formal employment, firm entry, and wages. I show that in Vietnam, these formal-sector outcomes capture only one quarter of the total non-farm income gains for households with limited access to formal employment. The remaining three quarters accrue through informal non-tradable enterprise activity that would be invisible in administrative data.

The paper develops a framework that connects the local multiplier mechanism to labor markets with skill-based segmentation and pervasive informality. When formal jobs are constrained, formal-sector expansion generates local demand for non-tradable services, raising returns to informal activities that are accessible to less-educated households. This framework generates predictions about sectoral reallocation, household labor diversification, and the composition of income gains that are consistent with the data. Less-educated households do not simply switch sectors; they temporarily diversify their labor portfolios across agriculture, formal, and informal work, with diversification tracing a sharp gradient across household types that maps onto the formal access constraint. In-migration of formal workers is part of the mechanism that generates local demand, but the welfare gains accrue to incumbent households through their informal enterprises.

These findings carry implications beyond the Vietnamese case. The informal demand channel requires two conditions: successful formal activity creation and a local informal sector flexible enough to absorb rising demand. Where both conditions are met, evaluations focused on formal outcomes will understate total benefits and mischaracterize distributional incidence. Where either is absent, the channel cannot operate. This distinction offers a potential explanation for why the place-based policy literature finds large effects in some settings and null effects in others, and it provides empirically testable conditions for predicting where zone programs are likely to generate inclusive welfare gains.

The analysis leaves open several questions for future work. The framework identifies conditions under which informal spillovers arise, but the longer-run dynamics remain unclear. Recent evidence suggests that informal enterprises can serve as stepping stones to formality over longer horizons, as labor supply shocks that initially increase informality eventually lead to greater formal firm entry and employment (Imbert & Ulyssea, 2026). Whether a similar dynamic operates in the context of place-based industrial policies, with informal enterprises that emerge in response to zone-driven demand eventually formalizing as zones mature, is an important question with direct implications for how the welfare gains documented here evolve over time. More broadly, the finding that informal channels account for the majority of welfare gains for less-educated households suggests that complementary investments in local service-economy capacity, small enterprise support, and connectivity to zone catchment areas may shape how broadly the benefits of place-based industrialization are shared.

## REFERENCES

- Abagna, M. A., Hornok, C., & Mulyukova, A. (2025). Place-based policies and household wealth in africa. *Journal of Development Economics*, 176, 103482.
- Balassa, B. (1964). The purchasing-power parity doctrine: A reappraisal. *Journal of Political Economy*, 72(6), 584–596.
- Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural africa: Concepts, dynamics, and policy implications. *Food Policy*, 26(4), 315–331.
- Bonnet, F., Vanek, J., Chen, M., et al. (2019). *Women and men in the informal economy: A statistical brief* (tech. rep. No. 1). International Labour Office, Geneva.
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.
- Busso, M., Gregory, J., & Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2), 897–947.
- Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230. <https://doi.org/10.1016/j.jeconom.2020.12.001>
- Criscuolo, C., Martin, R., Overman, H. G., & Van Reenen, J. (2019). Some causal effects of an industrial policy. *American Economic Review*, 109(1), 48–85.
- de Chaisemartin, C., Ciccia, D., D’Haultfoeuille, X., Knau, F., Malézieux, M., & Sow, D. (2024). *Event-study estimators and variance estimators computed by the did\_multiplegt\_dyn command* (Available at SSRN).
- de Chaisemartin, C., & d’Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- de Chaisemartin, C., & d’Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.
- Dercon, S. (2002). Income risk, coping strategies, and safety nets. *World Bank Research Observer*, 17(2), 141–166.

- Dix-Carneiro, R., Goldberg, P., Meghir, C., & Ulyssea, G. (2026). Trade and domestic distortions: The case of informality. *Econometrica*, 94(2), 573–618.
- Donovan, K., Lu, W. J., & Schoellman, T. (2023). Labor market dynamics and development. *Quarterly Journal of Economics*, 138(4), 2287–2325.
- Duranton, G., & Venables, A. J. (2018). *Place-based policies for development* (Working Paper No. 24562). National Bureau of Economic Research. <https://doi.org/10.3386/w24562>
- Fields, G. S. (2011). Labor market analysis for developing countries. *Labour Economics*, 18, S16–S22.
- Fox, M. L., & Gaal, M. S. (2008). Working out of poverty: Job creation and the quality of growth in africa.
- Gallé, J., Overbeck, D., Riedel, N., & Seidel, T. (2024). Place-based policies, structural change and female labor: Evidence from india’s special economic zones. *Journal of Public Economics*, 240, 105259. <https://doi.org/10.1016/j.jpubeco.2024.105259>
- Garin, A., & Rothbaum, J. (2025). The long-run impacts of public industrial investment on local development and economic mobility: Evidence from world war ii. *Quarterly Journal of Economics*, 140(1), 459–520.
- General Statistics Office of Vietnam and International Labor Organization. (2018). *2016 Report on informal employment in Vietnam* (tech. rep.). General Statistics Office of Vietnam and International Labor Organization.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Granata, J., Moroz, H. E., & Nguyen, N. T. (2023). *Identifying skills needs in Vietnam: The survey of detailed skills (english)*. <http://documents.worldbank.org/curated/en/099508509112311079>
- Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, 60(1), 126–142.
- Hornbeck, R., & Moretti, E. (2024). Estimating who benefits from productivity growth: Local and distant effects of city productivity growth on wages, rents, and inequality. *Review of Economics and Statistics*, 106(3), 587–607.
- Imbert, C., & Ulyssea, G. (2026). Rural migrants and urban informality: Evidence from brazil. *Econometrica*.

- Kline, P., & Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, 129(1), 275–331.
- La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3), 109–126.
- Lu, F., Sun, W., & Wu, J. (2023). Special economic zones and human capital investment: 30 years of evidence from china. *American Economic Journal: Economic Policy*, 15(3), 35–64.
- Lu, Y., Wang, J., & Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from China’s economic zone program. *American Economic Journal: Economic Policy*, 11(3), 325–360.
- Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7), 1159–1178.
- McCaig, B., & Pavcnik, N. (2018). Export markets and labor allocation in a low-income country. *American Economic Review*, 108(7), 1899–1941.
- McCaig, B., & Pavcnik, N. (2021). Entry and exit of informal firms and development. *IMF Economic Review*, 69, 540–575. <https://doi.org/10.1057/s41308-021-00142-8>
- Morduch, J. (1995). Income smoothing and consumption smoothing. *Journal of Economic Perspectives*, 9(3), 103–114.
- Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2), 373–377.
- Neumark, D., & Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics* (pp. 1197–1287, Vol. 5). Elsevier.
- Nguyen, T. T. A., Luu, M. D., & Trinh, D. C. (2016). The evolution of Vietnamese industry. In J. Page, F. Tarp, J. Rand, A. Shimeles, C. Newman, & M. Söderbom (Eds.), *Manufacturing transformation: Comparative studies of industrial development in Africa and emerging Asia*. Oxford University Press.
- Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Raymond, C. (2008). “No responsibility and no rice”: The rise and fall of agricultural collectivization in Vietnam. *Agricultural History*, 18(1), 43–61.
- Rosenzweig, M. R., & Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of Political Economy*, 97(4), 905–926.

- Rothenberg, A. D., Wang, Y., & Chari, A. (2025). When regional policies fail: An evaluation of indonesia's integrated economic development zones. *Journal of Development Economics*, 176, 103503.
- Samuelson, P. A. (1964). Theoretical notes on trade problems. *Review of Economics and Statistics*, 145–154.
- Tafese, T., Lay, J., & Tran, V. (2025). From fields to factories: Special economic zones, foreign direct investment, and labour markets in vietnam. *Journal of Development Economics*, 103467.
- Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of Political Economy*, 104(5), 1010–1046.
- Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, 12(1), 525–546.
- UNCTAD. (2019). World investment report 2019: Special economic zones. [https://unctad.org/system/files/official-document/wir2019\\_en.pdf](https://unctad.org/system/files/official-document/wir2019_en.pdf)
- United Nations Industrial Development Organization. (2019). Eco-industrial parks vietnam socio-economic requirements: A review of international and vietnamese experiences. <https://www.unido.org/sites/default/files/files/2019-05/3-Vietnam-Review-of-international-and-Vietnamese-experiences-1.pdf>
- Wang, J. (2013). The economic impact of special economic zones: Evidence from Chinese municipalities. *Journal of Development Economics*, 101, 133–147.
- World Bank Group. (2019). Vietnam development report 2019: Connecting vietnam for growth and shared prosperity. <https://documents1.worldbank.org/curated/en/590451578409008253/pdf/Vietnam-Development-Report-2019-Connecting-Vietnam-for-Growth-and-Shared-Prosperity.pdf>
- Zhao, C., & Qu, X. (2024). Place-based policies, rural employment, and intra-household resources allocation: Evidence from China's economic zones. *Journal of Development Economics*, 167, 103210.

# Supplementary Materials for Online Publication

Who Benefits from Place-Based Industrial Policies? Labor Market Adjustments and  
Household Welfare in Vietnam

Trinh Pham

KDI School of Public Policy and Management

tpham@kdischool.ac.kr

## CONTENTS

<b>A Theoretical Model</b>	<b>S1</b>
A.1 Model Setup . . . . .	S2
A.2 Comparative Statics: Effect of IZ Establishment . . . . .	S10
<b>B Appendix Tables and Figures</b>	<b>S21</b>
B.1 Appendix Tables . . . . .	S21
B.2 Appendix Figures . . . . .	S27
<b>C Distance-Based Analysis</b>	<b>S33</b>

## A THEORETICAL MODEL

In this section, I introduce a stylized local-economy model that clarifies how industrial zone establishment affects sectoral labor allocation, household diversification, and welfare when formal employment opportunities expand but remain imperfectly accessible. The model features two household types that differ in their access to formal jobs. Industrial zones raise the scale and productivity of the formal tradable sector, increasing formal wages. However, less-educated households face a binding barrier to formal participation—capturing rationing, screening, or hiring frictions—so their formal employment can expand with zone exposure but does not adjust freely. More-educated households, by contrast, have better access and are assumed to concentrate their labor in formal employment once the zone becomes sufficiently important.

The central mechanism operates through local demand spillovers. Higher formal earnings, especially among more-educated households, increase demand for locally produced non-tradable goods and services. Because these non-tradables are supplied locally—primarily through the informal activities of less-educated households—market clearing implies an increase in the equilibrium price of non-tradables. This price response raises the relative return to informal non-agricultural work compared with agriculture, inducing less-educated households to reallocate their remaining non-formal labor away from agriculture toward informal services. As a result, less-educated households respond to zone expansion along two margins: (i) a direct margin, as formal participation increases to the extent that access relaxes; and (ii) an indirect margin, as higher non-tradable prices shift the composition of non-formal labor toward informal services.

These forces imply a non-monotonic pattern for household labor diversification. When less-educated households are initially agriculture-heavy and formal access is limited, modest increases in formal work combined with a shift into informal services make labor portfolios more balanced across sectors, increasing diversification. With greater zone exposure, the informal service sector can absorb most non-formal labor, and diversification may flatten or decline as labor becomes concentrated in informal activities. The model also clarifies welfare implications: although higher non-tradable prices raise the cost of consumption, less-educated households are the primary suppliers of non-tradables and therefore benefit from the increase in informal returns. Consequently, industrial zone expansion generates positive income and welfare gains for less-educated house-

holds through both direct formal earnings and the local-demand channel, and it can produce heterogeneous effects across locations depending on local skill composition. These predictions provide a set of testable implications that guide the empirical analysis.

## A.1 Model Setup

**Endowment.** Consider a local economy that consists of a continuum of households indexed by skill type  $s \in \{H, L\}$ , where  $H$  represents households with skilled workers, and  $L$  presents households with unskilled workers.

Let  $M_H$  and  $M_L$  denote the mass of households of each type, with total mass  $M = M_H + M_L$ . Define the skill ratio (skill intensity) of the economy as:

$$\phi := \frac{M_H}{M_H + M_L} \in (0, 1) \quad (\text{A1})$$

Each household has one unit of labor to allocate across activities. This normalization can be interpreted as total household labor supply, with  $l_A$ ,  $l_F$ , and  $l_I$  representing shares of household labor in agriculture (A), formal non-agriculture (F), and informal non-agriculture (I), respectively.<sup>1</sup>

**Production.** Production in all three sectors uses labor as the sole variable input. Agriculture and formal non-agriculture produce tradable goods whose prices ( $p_A, p_F$ ) are determined in outside markets. Informal non-agriculture produces goods and services that are consumed locally at price  $p_I$ .

A household operating in agriculture or informal non-agriculture produces

$$\begin{aligned} y_A &= Z_A(l_A)^\alpha \\ y_I &= Z_I(l_I)^\alpha, \end{aligned} \quad (\text{A2})$$

where  $Z_A$  and  $Z_I$  denote agricultural and informal non-agricultural productivity, respectively, and  $\alpha \in (0, 1)$  captures decreasing returns to scale.<sup>2</sup>

---

<sup>1</sup>For simplicity, the model treats households as the decision making unit. The labor allocation decision represents the household' aggregate allocation across sectors, which can arise from a single member diversifying across activities (within-member diversification) or different members specializing in different sectors (across-member diversification).

<sup>2</sup>Assuming a common returns-to-scale parameter  $\alpha$  allows for closed-form solutions while preserving the key mechanisms.

The marginal revenue product of the two sectors would be:

$$\begin{aligned}\frac{\partial R_A}{\partial l_A} &= p_A Z_A \alpha (l_A)^{\alpha-1} \\ \frac{\partial R_I}{\partial l_I} &= p_I Z_I \alpha (l_I)^{\alpha-1}\end{aligned}\tag{A3}$$

The formal non-agricultural sector produces output according to:

$$y_F = \theta F(L^H, L^L),\tag{A4}$$

where  $L^H$  and  $L^L$  denote labor supplied by high- and low-skilled workers, respectively. For simplicity, the production technology is linear:

$$F(L^H, L^L) = \lambda L^H + (1 - \lambda)L^L,\tag{A5}$$

with  $\lambda \in (0, 1)$ . The formal good is tradable and sold at an exogenous price  $p_F$ .

Under perfect competition, firms hire labor until wages equal marginal products, yielding skill-dependent formal wages:

$$w_F^H = p_F \theta \lambda, \quad w_F^L = p_F \theta (1 - \lambda),\tag{A6}$$

where  $\theta > 0$  captures the scale and productivity of the formal non-agricultural sector and increases with industrial zone expansion.

**Assumption A1.** (*Skill Bias*): Formal sector productivity is skill-biased, with  $\lambda \in (0.5, 1)$ , implying that formal wages are higher for high-skilled than for low-skilled workers ( $w_F^H > w_F^L$ ).

**Preferences.** Households consume a bundle of tradable goods ( $C_T$ ) and non-tradable services ( $C_I$ ) with Cobb-Douglas utility:

$$U = C_T^{1-\mu} C_I^\mu\tag{A7}$$

where  $\mu \in (0, 1)$  is the expenditure share on non-tradables.

Given household income  $y$  and prices  $(1, p_I)$  where the tradable good is numeraire, optimal

consumption yields:

$$\begin{aligned} C_T &= (1 - \mu)y \\ C_I &= \frac{\mu y}{p_I} \end{aligned} \tag{A8}$$

Indirect utility is:

$$V(y, p_I) = \kappa \frac{y}{p_I^\mu} \tag{A9}$$

for some constant  $\kappa > 0$ . Given common equilibrium prices for non-tradables, indirect utility is proportional to real income, thus maximizing income is equivalent to maximizing utility.

**Households.** A type-S household chooses labor allocation  $(l_A^S, l_F^S, l_I^S)$  to maximize total income that is the sum of formal non-agricultural employment wage, and revenues from agricultural and informal non-farm household businesses:

$$\begin{aligned} \max_{l_F^S, l_A^S, l_I^S} \quad & w_F^S l_F^S + p_A Z_A (l_A^S)^\alpha + p_I Z_I (l_I^S)^\alpha \\ \text{s.t.} \quad & l_F^S + l_A^S + l_I^S = 1 \\ & l_F^S, l_A^S, l_I^S \geq 0 \end{aligned} \tag{A10}$$

**Assumption A2.** (*Formal Non-Agricultural Sector Constraint*): Assume that type-L households also face a constraint on formal sector participation

$$l_F^L \leq \bar{l}(\theta, \phi) \tag{A11}$$

where  $\bar{l}$  is increasing in  $\theta$  (e.g., more formal jobs available as the sector expands because of industrial zone establishment) but decreasing in  $\phi$  (e.g., L workers face more competition for formal jobs in areas with higher share of H workers). Specifically,

$$\bar{l}(\theta, \phi) := \frac{\bar{l}_0 \theta}{1 + \phi} < 1 \tag{A12}$$

for some  $\bar{l}_0 > 0$  and all  $\theta$  examined. This constraint captures labor market frictions that limit less-educated workers' access to formal employment, such as hiring barriers, skill requirements, or rationing.

This constraint is binding both before and after zone establishment, but relaxes as  $\theta$  increases:

$$\begin{aligned}\frac{\partial \bar{l}}{\partial \theta} &= \frac{\bar{l}_0}{1 + \phi} > 0 \\ \frac{\partial \bar{l}}{\partial \phi} &= -\frac{\bar{l}_0 \theta}{(1 + \phi)^2} < 0\end{aligned}\tag{A13}$$

Type-*H* households do not face such constraint.

Note that I interpret the constraint  $l_F^L \leq \bar{l}(\theta, \phi)$  as a reduced-form representation of rationing or screening in access to formal employment for type-*L* households. To understand the underlying micro foundations, let  $Vacancy(\theta)$  denote the measure of formal non-farm positions (“slots”) available in the local labor market. Since the economy has total mass  $M$  of households and each household supplies one unit of labor, it is natural to express formal job creation in per-household terms.

I assume that zone expansion increases formal labor demand proportionally, so that formal slots scale with the size of the economy:

$$Vacancy(\theta) = \bar{l}_0 \theta M$$

where  $\bar{l}_0 > 0$  is the number of formal positions created per household per unit of  $\theta$  (equivalently,  $Vacancy(\theta)/M = \bar{l}_0 \theta$  is the per-household vacancy rate). This formulation is consistent with a simple entry/capacity story in which the mass of formal firms is proportional to  $\theta M$  and each firm supplies a fixed number of formal positions.

Hiring is rationed via a weighted lottery or queue: type-*L* applicants have weight  $w_L = 1$  while type-*H* applicants have weight  $w_H = 2$ , capturing that higher-skill households are more likely to be selected because of screening, credentials, or networks. This parameterization is consistent with the empirical estimates in Section 5, where industrial zone establishment increases formal non-agricultural employment by roughly two to three times more for workers with a high school diploma than for those without (Table 4).<sup>3</sup>

<sup>3</sup>I view this as suggestive evidence for a relative advantage in formal sector access for high-skill households; the model uses  $w_H/w_L = 2$  as a parsimonious reduced-form representation rather than a directly identified structural parameter.

Then the probability (or feasible share) of formal employment for a type- $L$  household is

$$p_L(\theta, \phi) = \frac{\text{Vacancy}(\theta)}{w_L M_L + w_H M_H} = \frac{\bar{l}_0 \theta M}{(1 - \phi)M + 2\phi M} = \frac{\bar{l}_0 \theta}{1 + \phi} := \bar{l}(\theta, \phi)$$

so  $\theta$  increases access by expanding vacancies while  $\phi$  decreases access by intensifying effective competition.

I focus on parameter values such that type- $L$ 's desired formal participation exceeds  $\bar{l}(\theta, \phi)$  over the  $\theta$  range considered, so the constraint binds both before and after zone establishment.

### Type- $H$ Household Optimization.

**Assumption A3.** (*Type- $H$  Household Specialization*): For  $\theta \geq \theta_0$ , formal wages are sufficiently high that type- $H$  households specialize in formal employment:

$$\begin{aligned} l_F^H &= 1 \\ l_A^H = l_I^H &= 0 \end{aligned} \tag{A14}$$

Note that Assumption A3 is imposed for tractability. With the concave household-enterprise technologies  $y_j = Z_j(l_j)^\alpha$  and  $\alpha \in (0, 1)$ , marginal returns are unbounded as  $l_j \rightarrow 0$ , so a strict corner  $l_A^H = l_I^H = 0$  does not follow from finite wage comparisons alone.

Instead, higher formal wages make the optimal household-enterprise allocations arbitrarily small: any interior allocation satisfies  $w_F^H = p_j Z_j \alpha (l_j^H)^{\alpha-1} \Rightarrow l_j^H = (p_j Z_j \alpha / w_F^H)^{1/(1-\alpha)}$

Since  $w_F^H = p_F \lambda \theta$  increases in  $\theta$ , for  $\theta \geq \theta_0$  these non-formal allocations are negligible, and I approximate the resulting near-corner solution by full specialization  $l_F^H = 1$ .

Allowing interior solutions for type- $H$  households would not change the model's qualitative predictions for type- $L$  households—which is my primary focus, and would mainly affect magnitudes through aggregate income and non-tradable market clearing.

Type- $H$  household income would be:

$$y^H = w_F^H = p_F \lambda \theta \tag{A15}$$

**Type-L Household Optimization.** type- $L$  households face the binding constraints  $l_F^L = \bar{l}$  and optimally split remaining labor between agriculture and informal non-agricultural work. The remaining problem is

$$\max_{l_A^L, l_I^L} p_A Z_A (l_A^L)^\alpha + p_I Z_I (l_I^L)^\alpha \quad \text{s.t.} \quad l_A^L + l_I^L = 1 - \bar{l} := \tilde{l} \quad (\text{A16})$$

Let  $\lambda^*$  be the shadow price of non-formal labor, then first order conditions yield:

$$p_A Z_A \alpha (l_A^L)^{\alpha-1} = \lambda^* = p_I Z_I \alpha (l_I^L)^{\alpha-1} \quad (\text{A17})$$

**Definition A1.** (*Relative Attractiveness*): Let  $\omega$  denote the relative attractiveness of informal non-agriculture versus agricultural work:

$$\omega := \frac{l_I}{l_A} = \left( \frac{p_I Z_I}{p_A Z_A} \right)^{\frac{1}{1-\alpha}} \quad (\text{A18})$$

where the equality follows from Equation (A17). In other words,  $\omega$  measures the relative revenue potential of informal non-farm work versus agriculture.

- $\omega > 1$ : informal non-agricultural sector is more attractive, so  $l_I > l_A$  (informal non-agricultural heavy)
- $\omega < 1$ : agricultural sector is more attractive, so  $l_A > l_I$  (agriculture heavy)
- $\omega = 1$ : both sectors are equally attractive, so  $l_A = l_I$

For type- $L$  households, optimal labor allocation would be:

$$\begin{aligned} l_A^L &= \frac{1 - \bar{l}}{1 + \omega} \\ l_F^L &= \bar{l} \\ l_I^L &= \frac{\omega(1 - \bar{l})}{1 + \omega} \end{aligned} \quad (\text{A19})$$

**Lemma A1.** There exists a unique interior solution  $(l_A^L, l_I^L) \in (0, \tilde{l})^2$  to type- $L$  household problem.

*Proof.* Define  $G(l_I) := p_A Z_A \alpha (l_A^L)^{\alpha-1} - p_I Z_I \alpha (l_I^L)^{\alpha-1}$ .

As  $l_I \rightarrow 0^+ : G(l_I) \rightarrow -\infty$ .

As  $l_I \rightarrow \tilde{l}^- : G(l_I) \rightarrow +\infty$ .

By continuity, there exists  $l_I^* \in (0, \tilde{l})$  with  $G(l_I^*) = 0$ .

In addition,  $G'(l_I) = p_A Z_A \alpha (1 - \alpha) (l_A^L)^{\alpha-2} + p_I Z_I \alpha (1 - \alpha) (l_I^L)^{\alpha-2} > 0$ , so  $G$  is strictly increasing, and thus  $l_I^*$  is unique. ■

The income of a type- $L$  household would be:

$$y^L = w_F^L \bar{l} + p_A Z_A \left( \frac{1 - \bar{l}}{1 + \omega} \right)^\alpha + p_I Z_I \left( \frac{\omega(1 - \bar{l})}{1 + \omega} \right)^\alpha \quad (\text{A20})$$

Applying the first order condition (A17), this simplifies to:

$$y^L = p_F \theta (1 - \lambda) \bar{l} + p_A Z_A \frac{(1 - \bar{l})^\alpha}{(1 + \omega)^{\alpha-1}} \quad (\text{A21})$$

**Aggregate Output and Income.** As type- $H$  households specialize in formal non-farm sector, the total informal non-agricultural output of this economy is:

$$Y_I = M_L Z_I (l_I^L)^\alpha = M_L Z_I \left( \frac{\omega(1 - \bar{l})}{1 + \omega} \right)^\alpha \quad (\text{A22})$$

And the total income is:

$$Y^{total} = M_H p_F \theta \lambda + M_L \left( p_F \theta (1 - \lambda) \bar{l} + p_A Z_A \frac{(1 - \bar{l})^\alpha}{(1 + \omega)^{\alpha-1}} \right) \quad (\text{A23})$$

**Local Goods and Services Market.** With Cobb-Douglas preferences, the aggregate demand for non-tradables is:

$$D_I = \frac{\mu Y^{total}}{p_I} \quad (\text{A24})$$

**Characteristics of Equilibrium.** An equilibrium, given exogenous formal sector scale and productivity  $\theta$ , consists of labor allocations  $(l_A^S, l_F^S, l_I^S)$  where  $S \in \{H, L\}$ , and a price of non-tradable services  $p_I$  such that households optimally allocate labor given prices and constraints, and the local market for non-tradable services clears.

- *type-H household optimization*: more-educated households specialize in formal employment

$$(l_A^H, l_F^H, l_I^H) = (0, 1, 0)$$

- *type-L household optimization*: less-educated households face a binding constraint on formal employment  $l_F^L = \bar{l}(\theta, \phi)$ , and allocate remaining labor between agriculture and informal non-farm work such that:

- marginal returns are equalized:  $p_A Z_A \alpha (l_A^L)^{\alpha-1} = p_I Z_I \alpha (l_I^L)^{\alpha-1}$
- the time constraint holds:  $l_A^L + l_I^L = 1 - \bar{l}$

- *Local goods market clearing*: Aggregate demand for non-tradable services equals aggregate supply

$$D_I = Y_I \Rightarrow p_I = \frac{\mu Y^{total}}{M_L Z_I (l_I^L)^\alpha}$$

**Lemma A2.** For any  $\theta > 0$ , there exists a unique equilibrium  $p_I^* > 0$ .

*Proof.* Define the excess demand function as:

$$H(p_I) := D_I(p_I) - Y_I(p_I)$$

With Cobb–Douglas preferences, households spend a constant expenditure share  $\mu$  on informal services, so aggregate demand satisfies

$$D_I(p_I) = \frac{\mu Y^{total}(p_I)}{p_I}$$

Total informal output is given by:

$$Y_I(p_I) = M_L Z_I (l_I^L(p_I))^\alpha,$$

where  $l_I^L(p_I)$  is increasing in  $p_I$  and bounded above by  $1 - \bar{l}$ .

As  $p_I \rightarrow 0$ , nominal income from informal activities vanishes and total income converges to

income from tradable sectors, which is strictly positive. Aggregate demand therefore diverges:

$$\lim_{p_I \rightarrow 0} D_I(p_I) = +\infty, \quad \lim_{p_I \rightarrow 0} Y_I(p_I) = 0,$$

implying

$$\lim_{p_I \rightarrow 0} H(p_I) = +\infty$$

As  $p_I \rightarrow \infty$ , informal labor converges to its upper bound and informal output converges to a finite level  $\bar{Y}_I$ . Total income grows linearly in  $p_I$ , so aggregate demand converges to a finite constant:

$$\lim_{p_I \rightarrow \infty} D_I(p_I) = \mu \bar{Y}_I, \quad \lim_{p_I \rightarrow \infty} Y_I(p_I) = \bar{Y}_I$$

Since  $\mu \in (0, 1)$ ,

$$\lim_{p_I \rightarrow \infty} H(p_I) = -(1 - \mu)\bar{Y}_I < 0.$$

By continuity, there exists at least one equilibrium price  $p_I^* > 0$  such that  $H(p_I^*) = 0$ .<sup>4</sup>

Moreover, the excess demand function  $H(p_I)$  is strictly decreasing in  $p_I$ , implying that the equilibrium informal price  $p_I^*$  is unique. ■

Having established the existence and uniqueness of the equilibrium, I will now examine how equilibrium prices and labor allocations respond to changes in industrial zone exposure.

## A.2 Comparative Statics: Effect of IZ Establishment

In this setting, industrial zone establishment affects the local economy through changes in the scale and productivity of the formal non-agricultural sector, summarized by the parameter  $\theta$ . I examine how increases in  $\theta$  influence:

- Equilibrium prices (Proposition A1)
- Labor allocations (Proposition A2)

---

<sup>4</sup>The existence result relies on the boundedness of real demand for informal non-farm goods and services implied by homothetic preferences; Cobb–Douglas preferences provide a convenient closed-form representation.

- Household labor diversification (Proposition A3) and its heterogeneous responses across skill intensity levels under certain assumptions (Corollary A1)
- Household income (Proposition A4)

**Proposition A1.** (*Local Price Effect*). An increase in industrial zone exposure  $\theta$  raises the equilibrium price of non-tradable services  $p_I$ :

$$\frac{dp_I}{d\theta} > 0$$

*Proof.* The equilibrium informal price  $p_I$  is determined by clearing the local market for informal non-farm goods and services. Define the excess demand as:

$$\mathcal{H}(p_I, \theta) := D_I(p_I, \theta) - Y_I(p_I, \theta)$$

In equilibrium  $\mathcal{H}(p_I, \theta) = 0$ . By the implicit function theorem:

$$\frac{dp_I}{d\theta} = -\frac{\partial \mathcal{H} / \partial \theta}{\partial \mathcal{H} / \partial p_I} \quad (\text{A25})$$

I first consider the sign of  $\partial \mathcal{H} / \partial \theta$ . Holding  $p_I$  fixed, and increase in  $\theta$  affects both the supply and demand for non-tradables.

- Supply effect: higher  $\theta$  increases formal access for type- $L$  households, increasing formal non-farm labor allocation  $\bar{l}(\theta)$  and reducing labor for informal non-farm work:

$$\frac{\partial l_I^L}{\partial \theta} < 0$$

Since informal non-farm output is given by  $Y_I = Z_I(l_I^L)^\alpha$ , this implies

$$\frac{\partial Y_I}{\partial \theta} < 0$$

- Demand effect: an increase in  $\theta$  raises formal non-farm wages and total household income.

With Cobb-Douglas preferences, higher income increases expenditure on informal services:

$$\frac{\partial D_I}{\partial \theta} > 0$$

Combining these effects, an increase in  $\theta$  raises excess demand for informal non-farm goods and services at the initial price.

$$\frac{\partial \mathcal{H}}{\partial \theta} = \frac{\partial D_I}{\partial \theta} - \frac{\partial Y_I}{\partial \theta} > 0 \quad (\text{A26})$$

Next, I consider the sign of  $\partial \mathcal{H} / \partial p_I$ . Holding  $\theta$  fixed, consider an increase in the informal non-farm price  $p_I$ :

- Supply effect: A higher price  $p_I$  raises the return to informal non-farm work, inducing type- $L$  households to allocate labor toward these informal non-farm activities:

$$\frac{\partial l_I^L}{\partial p_I} > 0$$

This increases informal non-farm output:

$$\frac{\partial Y_I}{\partial p_I} > 0$$

- Demand effect: With Cobb-Douglas preferences, higher informal prices reduce real income and lower demand for informal services:

$$\frac{\partial D_I}{\partial p_I} < 0$$

Therefore the excess demand function is strictly decreasing in  $p_I$ :

$$\frac{\partial \mathcal{H}}{\partial p_I} = \frac{\partial D_I}{\partial p_I} - \frac{\partial Y_I}{\partial p_I} < 0 \quad (\text{A27})$$

Combining equations (A25), (A26) and (A27), it follows that  $\frac{dp_I}{d\theta} > 0$ . ■

Intuitively, industrial zone expansion affects non-tradable prices through two reinforcing mechanisms. On the supply side, expanded formal employment draws labor away from informal ser-

vices, reducing non-tradable supply. On the demand side, higher formal wages raise local income and increase demand for locally produced services. The resulting excess demand raises the equilibrium price of non-tradables,  $p_I$ , reflecting a local labor market analog of the Balassa–Samuelson effect (Balassa, 1964; Samuelson, 1964) and local spending multipliers (Moretti, 2010).

**Proposition A2.** (*Sectoral Reallocation*). *An increase in industrial zone exposure  $\theta$  leads to type-L households to reallocate non-formal labor away from agriculture and toward informal non-farm services.*

$$\frac{\partial l_F^L}{\partial \theta} > 0, \quad \frac{d}{d\theta} \left( \frac{l_I^L}{1-\bar{l}} \right) > 0 \quad (\text{A28})$$

*Proof.* For H-type households, they remain at corner solution for all  $\theta \geq \theta_0$  by Assumption A3.

For L-type households, they allocate more labor into the formal non-farm sector:

$$\frac{\partial l_F^L}{\partial \theta} = \frac{\partial \bar{l}}{\partial \theta} = \frac{\bar{l}_0}{1+\phi} > 0$$

The informal non-farm share of non-formal labor would be:

$$\frac{l_I^L}{1-\bar{l}} = \frac{\omega}{1+\omega}$$

Differentiating:

$$\frac{d}{d\theta} \left( \frac{\omega}{1+\omega} \right) = \frac{1}{(1+\omega)^2} \frac{d\omega}{d\theta}$$

Since  $\omega = \left( \frac{p_I Z_I}{p_A Z_A} \right)^{\frac{1}{1-\alpha}}$ :

$$\frac{\partial \omega}{\partial p_I} = \frac{\omega}{(1-\alpha)p_I} > 0$$

Therefore

$$\frac{d\omega}{d\theta} = \frac{\partial \omega}{\partial p_I} \frac{dp_I}{d\theta} = \frac{\omega}{(1-\alpha)p_I} \frac{dp_I}{d\theta} \quad (\text{A29})$$

has the same sign as  $\frac{dp_I}{d\theta}$ . Thus  $\frac{d}{d\theta} \left( \frac{l_I^L}{1-\bar{l}} \right) = \frac{d}{d\theta} \left( \frac{\omega}{1+\omega} \right) > 0$ . ■

Intuitively, higher equilibrium prices of informal service  $p_I$  raise the relative returns to informal non-farm work compared with agriculture. As a result, less-educated households substitute

within their non-formal labor allocation toward informal activities.

**Proposition A3.** (*Household Labor Diversification*). Define household labor diversification as

$$DIV^L := 1 - HHI^L, \quad HHI^L = \sum_{j \in \{A, F, I\}} (l_j^L)^2 \quad (\text{A30})$$

then with limited formal non-farm employment access, the effect of formal sector expansion on type-L household diversification is non-monotonic. Diversification increases when households are agriculture-heavy, and decreases when households are informal-non-farm-heavy.

*Proof.* Substitute the optimal allocations, the  $HHI^L$  index can be decomposed into a formal-employment component and a within-non-formal allocation component:

$$\begin{aligned} HHI^L &= \left( \frac{1 - \bar{l}}{1 + \omega} \right)^2 + \bar{l}^2 + \left( \frac{\omega(1 - \bar{l})}{1 + \omega} \right)^2 \\ &= \bar{l}^2 + (1 - \bar{l})^2 \frac{1 + \omega^2}{(1 + \omega)^2} \\ &= \bar{l}^2 + (1 - \bar{l})^2 HHI_{A,I} \end{aligned} \quad (\text{A31})$$

The total derivative is:

$$\begin{aligned} \frac{dHHI^L}{d\theta} &= 2\bar{l} \frac{d\bar{l}}{d\theta} - 2(1 - \bar{l}) \frac{d\bar{l}}{d\theta} HHI_{A,I} + (1 - \bar{l})^2 \frac{\partial HHI_{A,I}}{\partial \omega} \frac{d\omega}{d\theta} \\ &= 2 \left[ \bar{l} - (1 - \bar{l}) HHI_{A,I} \right] \frac{d\bar{l}}{d\theta} + (1 - \bar{l})^2 \frac{2(\omega - 1)}{(1 + \omega)^3} \frac{d\omega}{d\theta} \end{aligned} \quad (\text{A32})$$

Because  $DIV^L := 1 - HHI^L$ :

$$\frac{dDIV^L}{d\theta} = \underbrace{-2 \left[ \bar{l} - (1 - \bar{l}) HHI_{A,I} \right] \frac{d\bar{l}}{d\theta}}_{\text{formal non-farm expansion effect}} \quad \underbrace{-(1 - \bar{l})^2 \frac{2(\omega - 1)}{(1 + \omega)^3} \frac{d\omega}{d\theta}}_{\text{informal price-induced reallocation effect}} \quad (\text{A33})$$

The sign of  $[\bar{l} - (1 - \bar{l}) HHI_{A,I}]$  determines whether formal non-farm expansion increases or decreases diversification.

Note that  $HHI_{A,I} \in [0.5, 1]$ . At  $\omega = 1$ ,  $HHI_{A,I} = 0.5$  then

$$[\bar{l} - (1 - \bar{l}) HHI_{A,I}] = \frac{3\bar{l} - 1}{2}$$

which is positive if  $\bar{l} > 1/3$  (i.e., formal non-farm expansion reduces diversification), and negative if  $\bar{l} < 1/3$  (i.e., formal non-farm expansion increases diversification).

Intuitively, maximum diversification occurs at  $l_A = l_F = l_I = 1/3$ . If  $\bar{l} < 1/3$ , households are under-represented in formal non-farm sector, so expanding  $\bar{l}$  moves toward increased diversification. If  $\bar{l} > 1/3$ , households are over-represented in formal non-farm sector, so expansion further concentrates household labor in this sector.

Combining with Equation (A29), Proposition A1 and Assumption A2, sufficient conditions would be:

1. Diversification increases  $\left(\frac{dDIV^L}{d\theta} > 0\right)$  when  $\omega < 1$  and  $\bar{l}$  is not too large.
2. Diversification decreases  $\left(\frac{dDIV^L}{d\theta} < 0\right)$  when  $\omega > 1$  and/or  $\bar{l}$  is large.

■

Intuitively, formal non-farm sector expansion affects diversification through two channels. First, as more type- $L$  households gain access to formal non-farm employment, they spread labor across more sectors, increasing diversification when formal access is initially limited. Second, rising informal non-farm prices induce households to rebalance non-formal labor away from agriculture toward informal non-farm work. At low levels of industrial zone exposure, both effects increase diversification. As exposure increases further, informal non-farm activities absorb most non-formal labor, and the price effect reverses—further shifts toward informal reduce diversification. With limited formal access, the first effect remains positive but small, while the price effect dominates the dynamics, generating the non-monotonic pattern.

**Lemma A3.** (Cross-partial decomposition). Assume stability  $\mathcal{H}_{p_I}(p_I, \theta, \phi) < 0$  and  $\mathcal{H}_\theta(p_I, \theta, \phi) > 0$  in the region of interest. Then the equilibrium price response satisfies

$$\frac{\partial}{\partial \phi} \left( \frac{dp_I}{d\theta} \right) = -\frac{\mathcal{H}_{\theta\phi}\mathcal{H}_{p_I} - \mathcal{H}_\theta\mathcal{H}_{p_I\phi}}{\mathcal{H}_{p_I}^2} = \frac{dp_I}{d\theta} \left( \frac{\mathcal{H}_{\theta\phi}}{\mathcal{H}_\theta} - \frac{\mathcal{H}_{p_I\phi}}{\mathcal{H}_{p_I}} \right). \quad (\text{A34})$$

In particular,

$$\frac{\partial}{\partial \phi} \left( \frac{dp_I}{d\theta} \right) > 0 \iff \frac{\mathcal{H}_{\theta\phi}}{\mathcal{H}_\theta} > \frac{\mathcal{H}_{p_I\phi}}{\mathcal{H}_{p_I}}$$

*Proof.* From equilibrium  $\mathcal{H}(p_I, \theta, \phi) = 0$  and stability  $\mathcal{H}_{p_I} < 0$ , the implicit function theorem yields

$$\frac{dp_I}{d\theta} = -\frac{\mathcal{H}_\theta}{\mathcal{H}_{p_I}}$$

Differentiating with respect to  $\phi$  gives

$$\frac{\partial}{\partial \phi} \left( \frac{dp_I}{d\theta} \right) = -\frac{\mathcal{H}_{\theta\phi}\mathcal{H}_{p_I} - \mathcal{H}_\theta\mathcal{H}_{p_I\phi}}{\mathcal{H}_{p_I}^2}$$

Using  $\frac{dp_I}{d\theta} = -\frac{\mathcal{H}_\theta}{\mathcal{H}_{p_I}}$  and rearranging yields (A34). ■

**Assumption A4.** (*Composition-dominance condition*). In the region of interest, district skill intensity  $\phi$  amplifies the  $\theta$ -induced shift in excess demand for non-tradables more than it changes the price sensitivity of excess demand:

$$\frac{\mathcal{H}_{\theta\phi}}{\mathcal{H}_\theta} > \frac{\mathcal{H}_{p_I\phi}}{\mathcal{H}_{p_I}}$$

**Lemma A4.** (*Amplified price response under composition dominance*). Under Assumptions A2–A3, stability ( $\mathcal{H}_{p_I} < 0$ ), and Assumption A4, the equilibrium price response to formal sector expansion is increasing in district skill intensity:

$$\frac{\partial}{\partial \phi} \left( \frac{dp_I}{d\theta} \right) > 0$$

*Proof.* By Proposition A1,  $dp_I/d\theta > 0$ . Lemma A3 implies

$$\frac{\partial}{\partial \phi} \left( \frac{dp_I}{d\theta} \right) = \frac{dp_I}{d\theta} \left( \frac{\mathcal{H}_{\theta\phi}}{\mathcal{H}_\theta} - \frac{\mathcal{H}_{p_I\phi}}{\mathcal{H}_{p_I}} \right)$$

Assumption A4 makes the bracketed term strictly positive, hence  $\frac{\partial}{\partial \phi} \left( \frac{dp_I}{d\theta} \right) > 0$ . ■

**Corollary A1.** (*Diversification and Skill Ratio*). Fix  $(\theta, \phi)$  in the agriculture-heavy region  $\omega(\theta, \phi) < 1$  and with sufficiently limited formal access  $\bar{l}(\theta, \phi)$  such that Proposition A3 implies  $dDIV^L/d\theta > 0$ .

Under Assumption A4 (so that  $\partial_\phi(dp_I/d\theta) > 0$ ), the diversification response to formal sector expansion is stronger in more skill-intensive districts whenever the  $\phi$ -amplification of the price-induced reallocation channel dominates the  $\phi$ -attenuation of the direct formal-access channel, i.e.

$$\frac{\partial}{\partial \phi} \Delta_P(\theta, \phi) > -\frac{\partial}{\partial \phi} \Delta_F(\theta, \phi), \quad (\text{A35})$$

where

$$\Delta_F(\theta, \phi) := -2\left[\bar{l} - (1 - \bar{l})HHI_{A,I}(\omega)\right] \frac{\partial \bar{l}}{\partial \theta}, \quad \Delta_P(\theta, \phi) := -(1 - \bar{l})^2 \frac{2(\omega - 1)}{(1 + \omega)^3} \frac{d\omega}{d\theta}.$$

In particular, under (A35),

$$\frac{\partial}{\partial \phi} \left( \frac{dDIV^L}{d\theta} \right) > 0$$

*Proof.* From Proposition A3, write the diversification response as

$$\frac{dDIV^L}{d\theta} = \Delta_F(\theta, \phi) + \Delta_P(\theta, \phi)$$

Differentiating with respect to  $\phi$  yields

$$\frac{\partial}{\partial \phi} \left( \frac{dDIV^L}{d\theta} \right) = \frac{\partial \Delta_F}{\partial \phi} + \frac{\partial \Delta_P}{\partial \phi}$$

Under condition (A35), we have

$$\frac{\partial \Delta_P}{\partial \phi} > -\frac{\partial \Delta_F}{\partial \phi}$$

and therefore

$$\frac{\partial}{\partial \phi} \left( \frac{dDIV^L}{d\theta} \right) > 0$$

which proves the corollary. ■

The corollary highlights that heterogeneity in diversification responses is primarily driven by the local-demand/price channel. Districts with higher skill share  $\phi$  experience larger formal income gains when industrial zones expand, which increases demand for non-tradable services and raises informal returns more strongly under composition dominance. Because low-skill households in high- $\phi$  districts also spend a larger share of time in non-formal activities (due to tighter formal access), the induced change in informal returns generates a larger reallocation away from agriculture and, in agriculture-heavy areas, a stronger increase in diversification. I test these predictions by examining whether non-tradable return proxies, low-skill sectoral reallocation, and low-skill diversification respond more strongly to zone establishment in districts with higher  $\phi$ , particularly in locations that are agriculture-heavy prior to zone expansion.

**Proposition A4.** (*Income Gains for type-L Households*). *In equilibrium, industrial zone expansion increases the income of type-L households, despite their limited access to formal employment.*

*Proof.* Type-L household income is

$$y^L = w_F^L(\theta) \bar{l}(\theta, \phi) + p_A Z_A (l_A^L)^\alpha + p_I(\theta) Z_I (l_I^L)^\alpha$$

with  $l_F^L = \bar{l}(\theta, \phi)$  and  $l_A^L + l_I^L = 1 - \bar{l}(\theta, \phi)$ .

Taking the total derivative with respect to  $\theta$ ,

$$\frac{dy^L}{d\theta} = \frac{d}{d\theta} \left( w_F^L \bar{l} \right) + p_A Z_A \alpha (l_A^L)^{\alpha-1} \frac{dl_A^L}{d\theta} + p_I Z_I \alpha (l_I^L)^{\alpha-1} \frac{dl_I^L}{d\theta} + Z_I (l_I^L)^\alpha \frac{dp_I}{d\theta}.$$

By the type-L first-order condition for interior non-formal allocations,

$$p_A Z_A \alpha (l_A^L)^{\alpha-1} = p_I Z_I \alpha (l_I^L)^{\alpha-1} := \lambda^*$$

and using  $dl_A^L/d\theta + dl_I^L/d\theta = -d\bar{l}/d\theta$ , the reallocation terms simplify to

$$\lambda^* \left( \frac{dl_A^L}{d\theta} + \frac{dl_I^L}{d\theta} \right) = -\lambda^* \frac{d\bar{l}}{d\theta}$$

Expanding  $\frac{d}{d\theta} (w_F^L \bar{l}) = \bar{l} \frac{dw_F^L}{d\theta} + w_F^L \frac{d\bar{l}}{d\theta}$  yields the decomposition

$$\frac{dy^L}{d\theta} = \bar{l} \frac{dw_F^L}{d\theta} + (w_F^L - \lambda^*) \frac{d\bar{l}}{d\theta} + Z_I (l_I^L)^\alpha \frac{dp_I}{d\theta}. \quad (\text{A36})$$

Each term in (A36) is positive:  $dw_F^L/d\theta > 0$  and  $\bar{l} > 0$ ; the formal constraint binds so  $w_F^L > \lambda^*$  and  $d\bar{l}/d\theta > 0$ ; and Proposition A1 implies  $dp_I/d\theta > 0$ . Hence  $dy^L/d\theta > 0$ . ■

Equation (A36) highlights two sources of type-L income gains: direct formal-sector gains (higher formal wages and a relaxation of the binding access constraint) and an indirect local-demand channel that raises the return to informal activities through a higher non-tradable price  $p_I$ . The next corollary shows that these nominal gains translate into welfare gains despite the increase in  $p_I$ , because type-L households are net sellers of non-tradables in equilibrium.

**Corollary A2.** (Welfare gains for type-L households). Under Assumptions A2–A3 and Proposition A1, type-L welfare increases with industrial zone exposure:

$$\frac{dV^L}{d\theta} > 0$$

*Proof.* Since  $V^L = \kappa y^L / p_I^\mu$ ,

$$\frac{dV^L}{d\theta} = \frac{\kappa}{p_I^\mu} \left( \frac{dy^L}{d\theta} - \mu \frac{y^L}{p_I} \frac{dp_I}{d\theta} \right)$$

Substitute (A36):

$$\frac{dV^L}{d\theta} = \frac{\kappa}{p_I^\mu} \left[ \bar{l} \frac{dw_F^L}{d\theta} + (w_F^L - \lambda^*) \frac{d\bar{l}}{d\theta} + \left( Z_I (l_I^L)^\alpha - \mu \frac{y^L}{p_I} \right) \frac{dp_I}{d\theta} \right] \quad (\text{A37})$$

By market clearing in the non-tradable sector,

$$M_L Z_I (l_I^L)^\alpha = D_I = \frac{\mu Y^{total}}{p_I} = \frac{\mu}{p_I} (M_L y^L + M_H y^H)$$

so

$$Z_I (l_I^L)^\alpha - \mu \frac{y^L}{p_I} = \frac{\mu}{p_I} \frac{M_H}{M_L} y^H > 0 \quad (\phi \in (0, 1))$$

All remaining terms in (A37) are positive and Proposition A1 implies  $dp_I/d\theta > 0$ , hence  $dV^L/d\theta > 0$ . ■

*Discussion: Implications for Income (Non-)Divergence.* Type-H households specialize in formal employment, so  $y^H(\theta) = p_F \lambda \theta$  and  $dy^H/d\theta = p_F \lambda$ . For type-L households,

$$\frac{dy^L}{d\theta} = \underbrace{\bar{l} \frac{dw_F^L}{d\theta} + (w_F^L - \lambda^*) \frac{d\bar{l}}{d\theta}}_{=: G_F^L(\theta, \phi)} + \underbrace{Z_I (l_I^L)^\alpha \frac{dp_I}{d\theta}}_{=: G_I^L(\theta, \phi)}$$

Thus, comparability of income gains (e.g.  $dy^L/d\theta \geq dy^H/d\theta$ ) holds whenever

$$G_I^L(\theta, \phi) \geq p_F \lambda - G_F^L(\theta, \phi). \quad (\text{A38})$$

The term  $G_F^L(\theta, \phi) \geq 0$  captures type-L's direct formal-sector gains (higher formal wages on

existing formal work and the benefit of relaxing the binding access constraint). Equation (A38) therefore implies that local-demand spillovers need only fill the *remaining* gap relative to type- $H$ 's formal-income gains.

A useful benchmark is the limiting case in which direct formal gains for type- $L$  are small (e.g., low baseline formal access and limited access expansion), in which case (A38) is well-approximated by  $Z_I(l_I^L)^\alpha \frac{dp_I}{d\theta} \approx p_F \lambda$ .

## B APPENDIX TABLES AND FIGURES

### B.1 Appendix Tables

Table B1: Establishment of Industrial Zones Across Districts

	Share of districts hosting zone (1)	Share of districts within 15-km radius of zone (2)
2004-2008	0.102	0.549
2009-2013	0.033	0.060
2014-2020	0.010	0.023
Never-treated	0.855	0.367

*Notes:* This table presents the percentage of districts that either have an industrial zone within their boundaries (Column 1) or are located within a 15-kilometer radius of one (Column 2).

Table B2: Industrial Zones and Labor Participation, Ages 25–64

	All Individuals (1)	Without High School Diploma (2)	With High School Diploma (3)
post-ATT	0.002 (0.004)	0.000 (0.004)	0.012 (0.008)
District FE	Y	Y	Y
Year FE	Y	Y	Y
Mean Outcome	0.92	0.92	0.93
p-value pre-trend	0.513	0.420	0.772
N(Individual-Period)	300319	243160	56950
N(District Switcher-Period)	839	839	826

*Notes:* The sample includes individuals aged 25–64. Outcome is measured as whether an individual engaged in any economic activities during the last 12 months. Estimates are derived using the method proposed by de Chaisemartin and d'Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Data from VHLSS 2004-2020. Sampling weights are applied throughout.

Table B3: Industrial Zones and Working Members

	Without High School Graduate (1)	With High School Graduate (2)
post-ATT	0.016 (0.017)	0.008 (0.026)
District FE	Y	Y
Year FE	Y	Y
Mean Outcome	1.71	1.98
p-value pre-trend	0.807	0.840
N(Household-Period)	96363	54657
N(District Switcher-Period)	839	832

*Notes:* The outcome is number of household members aged 25–64 working during the past 12 months. Estimates are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Data from VHLSS 2004-2020. Sampling weights are applied throughout.

Table B4: Industrial Zones and Household Labor Diversification

## Heterogeneity by District-level Skill Intensity

Sample includes households without any members ever completing high school education

	Diversification Index		Across-Member		Within-Member	
	Low-Skill (1)	High-Skill (2)	Low-Skill (3)	High-Skill (4)	Low-Skill (5)	High-Skill (6)
Period1-ATT	-0.005 ( 0.005)	0.004 ( 0.014)	-0.020 ( 0.011)	0.021 ( 0.031)	-0.010 ( 0.012)	0.005 ( 0.032)
Period4-ATT	0.020 ( 0.008)	0.060 ( 0.030)	0.055 ( 0.022)	0.129 ( 0.069)	0.039 ( 0.022)	0.127 ( 0.054)
Period6-ATT	0.001 ( 0.012)	0.036 ( 0.040)	0.020 ( 0.028)	0.056 ( 0.091)	0.020 ( 0.027)	0.078 ( 0.062)
post-ATT	0.009 ( 0.007)	0.030 ( 0.025)	0.027 ( 0.017)	0.064 ( 0.058)	0.022 ( 0.017)	0.061 ( 0.042)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.18	0.20	0.48	0.53	0.35	0.38
p-value pre-trend	0.944	0.277	0.733	0.534	0.577	0.386
N(Household-Period)	80862	9099	80862	9099	80862	9099
N(District Switcher-Period)	675	140	675	140	675	140

*Notes:* This table shows the effects of industrial zone exposure on household-level labor diversification strategies across districts with different level of skilled labor demand. In Columns (1)–(2), the dependent variable calculated as 1 minus the Herfindahl–Hirschman Index (HHI) using the share of household labor hours in each sector (higher values indicate greater diversification). In Columns (3)–(4), the dependent variable is whether the household has different members working in different sectors. In Columns (5)–(6), the dependent variable is whether the household has at least one member working in multiple sectors. Columns (1), (3), and (5) include “low-skill” districts only; and Columns (2), (4), and (6) include “high-skill” districts only. A district is classified as “Low-Skill” if its share of working individuals with a high school diploma or a post-secondary degree is below the 75<sup>th</sup> percentile of the district-level distribution in 2002, based on data from the 2002 VHLSS; districts above this threshold are classified as “High-Skill.” Estimates are derived using de Chaisemartin et al. (2024)’s method. Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated for pre-treatment periods. Sampling weights are applied throughout.

Table B5: Industrial Zones and Informal Non-Tradable Non-Farm Business Performance

Heterogeneity by District-level Skill Intensity

	Business (0/1)		Revenue		Profits	
	Low-Skill	High-Skill	Low-Skill	High-Skill	Low-Skill	High-Skill
	(1)	(2)	(3)	(4)	(5)	(6)
post-ATT	0.026 (0.012)	0.084 (0.037)	0.979 (0.698)	3.546 (3.529)	0.735 (0.403)	1.732 (2.011)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.23	0.29	7.49	12.45	4.51	7.31
p-value pre-trend	0.348	0.958	0.804	0.247	0.483	0.659
N(Household-Period)	83974	9779	83974	9779	83974	9779
N(District Switcher-Period)	675	140	675	140	675	140

*Notes:* This table shows the effects of industrial zone exposure on informal business activity among less-educated households for non-tradable industries, across districts with different level of skilled labor demand. Non-tradable industries include construction, retail, wholesale, food services, accommodation, transport, personal services, repair services, healthcare, and education. Tradable industries include manufacturing, mining, and other activities producing goods for external markets. Columns (1), (3), and (5) include “low-skill” districts only; and Columns (2), (4), and (6) include “high-skill” districts only. A district is classified as “Low-Skill” if its share of working individuals with a high school diploma or a post-secondary degree is below the 75<sup>th</sup> percentile of the district-level distribution in 2002, based on data from the 2002 VHLSS; districts above this threshold are classified as “High-Skill.” Profits and revenue are winsorized at the 5th and 95th percentiles within each survey year, education group, and treatment status, conditional on positive values. Estimates of equation (7) are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Sampling weights are applied throughout.

Table B6: Industrial Zones, Household Consumption, and Human Capital Investment

Sample includes households without any members completing high school education

	Food Consumption (1)	Non-food Consumption (2)	Human Capital Investment (3)	School Enrollment (4)
post-ATT	0.007 (0.010)	0.011 (0.004)	0.006 (0.002)	0.020 (0.009)
District FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Mean Outcome	0.31	0.11	0.04	0.81
p-value pre-trend	0.245	0.182	0.946	0.880
N(Household/Individual-Period)	19239	19239	19239	94169
N(District Switcher-Period)	827	827	827	839

*Notes:* This table shows the effects of industrial zone exposure on welfare measures among less-educated households. In Column (1), the outcome is food consumption expenditure, combining regular food and drink items consumed in the past 30 days and annualized (rice, staple grains, meat, fish, seafood, eggs, tofu, vegetables, beans, fruits, cooking oil, sugar, coffee, tea, alcohol, tobacco, and outdoor meals) and food expenditure during major holidays over the past 12 months. In Column (2), the outcome is non-food consumption expenditure, comprising daily recurring non-food items (fuel, personal care products, parking fees, haircuts) reported over the past 30 days and annualized, and annual non-food items (clothing, footwear, household goods, appliances, communications, entertainment, travel, domestic services). This measure excludes food, housing, education, and health expenditure. Both food and non-food measures include purchased and self-supplied items. In Column (3), the outcome is total spending on education (tuition, textbooks, uniforms, tutoring) and health (in-patient and out-patient treatment). In Column (4), the outcome is whether a household member aged 10–18 attended school in the past 12 months. All monetary values are expressed in monthly per capita terms, in 2010 million VND. Estimates are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Standard errors, clustered at the district level, are shown in parentheses. Mean outcomes are calculated based on pre-treatment periods. Data from VHLSS 2004–2020. Sampling weights are applied throughout.

Table B7: Robustness Checks: Industrial Zones and Selected Outcomes

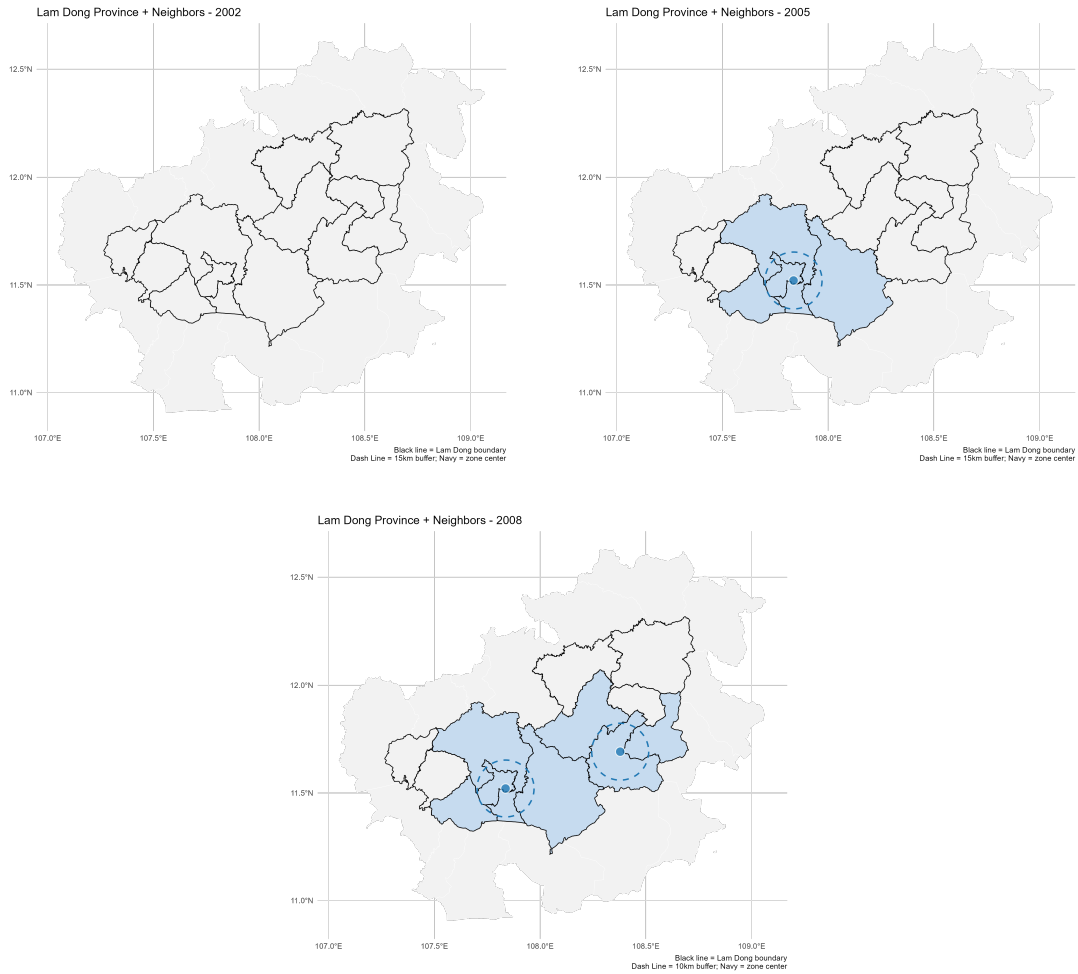
Treatment Effect of Period-4 Post-Zone Establishment

	Sectoral Employment		Labor Diversification		Income and Business	
	Agriculture	Informal Non-Ag	Index	Within Member	Informal Income	Non-Tradable Business (0/1)
	(1)	(2)	(3)	(4)	(5)	(6)
1. Baseline	-0.073 (0.017)	0.056 (0.016)	0.026 (0.008)	0.056 (0.020)	0.335 (0.065)	0.046 (0.015)
2. Including Not-Yet-Treated	-0.068 (0.017)	0.052 (0.015)	0.026 (0.008)	0.057 (0.020)	0.322 (0.065)	0.043 (0.015)
3. Province-level Clustering	-0.073 (0.019)	0.056 (0.017)	0.026 (0.010)	0.056 (0.026)	0.335 (0.061)	0.046 (0.015)
4. Callaway and Sant'Anna (2021)	-0.067 (0.017)	0.049 (0.016)	0.017 (0.008)	0.034 (0.021)	0.329 (0.067)	0.041 (0.016)
5. TWFE Event-Study	-0.032 (0.014)	0.047 (0.013)	0.020 (0.007)	0.037 (0.016)	0.171 (0.057)	0.031 (0.011)
6. Adding Demographic Controls	-0.073 (0.017)	0.056 (0.016)	0.025 (0.008)	0.056 (0.020)	0.341 (0.066)	0.048 (0.015)
7. Restricting to Registered Residents	-0.082 (0.018)	0.063 (0.017)	0.029 (0.009)	0.067 (0.023)	0.305 (0.074)	0.051 (0.018)
8. Timing Mismatch	-0.069 (0.016)	0.053 (0.015)	0.025 (0.008)	0.057 (0.020)	0.331 (0.065)	0.045 (0.016)
9. Excluding 15–25km Districts	-0.108 (0.019)	0.086 (0.017)	0.034 (0.009)	0.083 (0.024)	0.431 (0.073)	0.061 (0.020)
10. Household 3-Wave Panel	-0.043 (0.048)	0.051 (0.046)	0.025 (0.018)	0.058 (0.038)	0.324 (0.121)	0.044 (0.035)
11. Province Analysis	-0.105 (0.065)	0.075 (0.054)	0.014 (0.018)	0.029 (0.071)	0.389 (0.279)	0.102 (0.125)

Notes: Each cell reports the estimated effect at period 4 (approximately 8 years after zone establishment). Row 1 (Baseline) estimates equation (7) using the de Chaisemartin and d'Haultfoeuille (2024) estimator with never-treated districts as the comparison group and standard errors clustered at the district level. Row 2 extends the comparison group to include not-yet-treated districts. Row 3 clusters standard errors at the province level. Row 4 uses the alternative staggered difference-in-differences estimator of Callaway and Sant'Anna (2021). Row 5 reports estimates from a conventional two-way fixed effects event-study specification. Row 6 includes individual-level demographic controls (age, gender, and ethnic minority status). Row 7 restricts the sample to individuals with long-term permanent household registration in the same commune (for household-level analysis, this means household head having household registration in the same commune). Row 8 addresses the mismatch between annual zone establishment and biennial survey data by restricting to districts where treatment status changed between survey waves, following the split-sample strategy of de Chaisemartin and d'Haultfoeuille (2024). Row 9 excludes districts 15–25 km from any industrial zone to address potential spillover contamination of the comparison group. In Row 10, estimates are based on a household panel constructed by pooling four three-wave panels (2004–2006–2008, 2010–2012–2014, 2012–2014–2016, and 2014–2016–2018). When a household-year observation appears in multiple panels, it is included only once to avoid double-counting. The year 2020 is excluded because it is drawn from a different census-based sampling frame and cannot be linked to previous surveys. Row 11 reports province-level estimates using the cumulative number of zones as a continuous treatment variable, following de Chaisemartin and d'Haultfoeuille (2024), with standard errors clustered at the province level. Columns (1)–(2) report sectoral employment shares among individuals aged 25–64 without a high school diploma. Columns (3)–(4) report household labor diversification measures for households without a high school graduate. Columns (5)–(6) report household-level informal non-agricultural income and non-tradable business formation for households without a high school graduate. Standard errors, clustered at the district level unless otherwise noted, are shown in parentheses. Sampling weights are applied throughout.

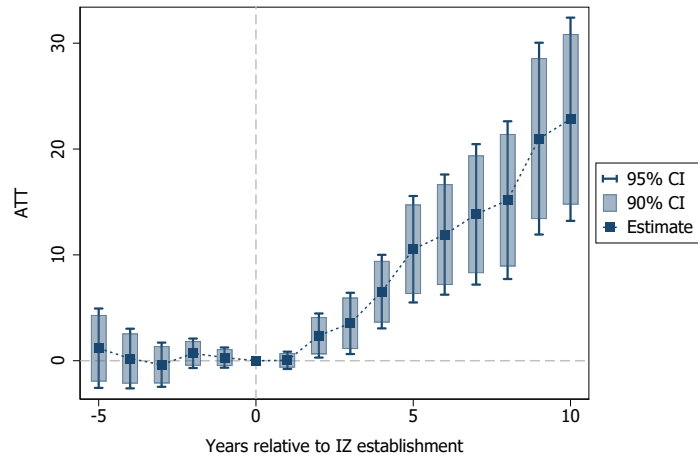
## B.2 Appendix Figures

Figure B1: Mapping Zones and Identification of District Treatment Status



*Notes:* The figure illustrates the staggered treatment status of districts over time. Districts shaded in blue are classified as treated starting from the year they first fall within a pre-specified buffer of an industrial zone geographic centroid. Once a district is treated, it remains in the treatment group in all subsequent years. Districts shaded in grey are either not yet treated or never treated during the study period. Treatment timing varies across districts depending on when nearby zones become active.

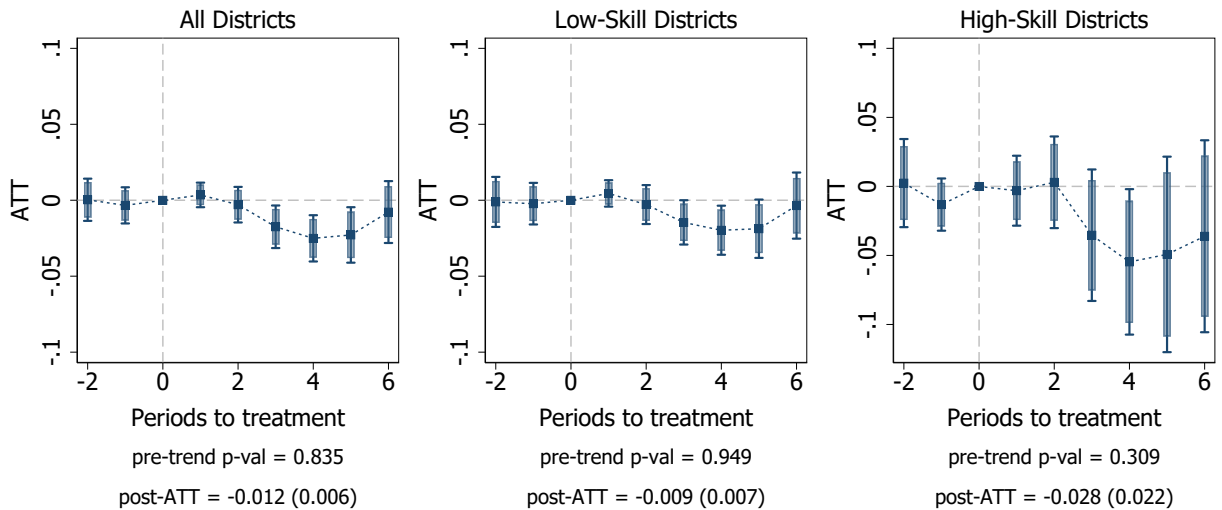
Figure B2: Industrial Zones and Manufacturing Firms Creation



Notes: This figure shows the effects of industrial zone establishment on the number of manufacturing firms per district, using data from the Vietnam Enterprise Survey 2001–2018. Treatment is defined as falling within a 10 km radius of a zone centroid. Square markers indicate the point estimates of the coefficients. Darker vertical lines with caps show 95% confidence intervals, and lighter bars represent 90% confidence intervals. Estimates are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024), with district and year fixed effects, and baseline district characteristics interacted with linear time trends. Standard errors clustered at the district level.

Figure B3: Industrial Zones and Distance to a Balanced Labor Portfolio Across Sectors

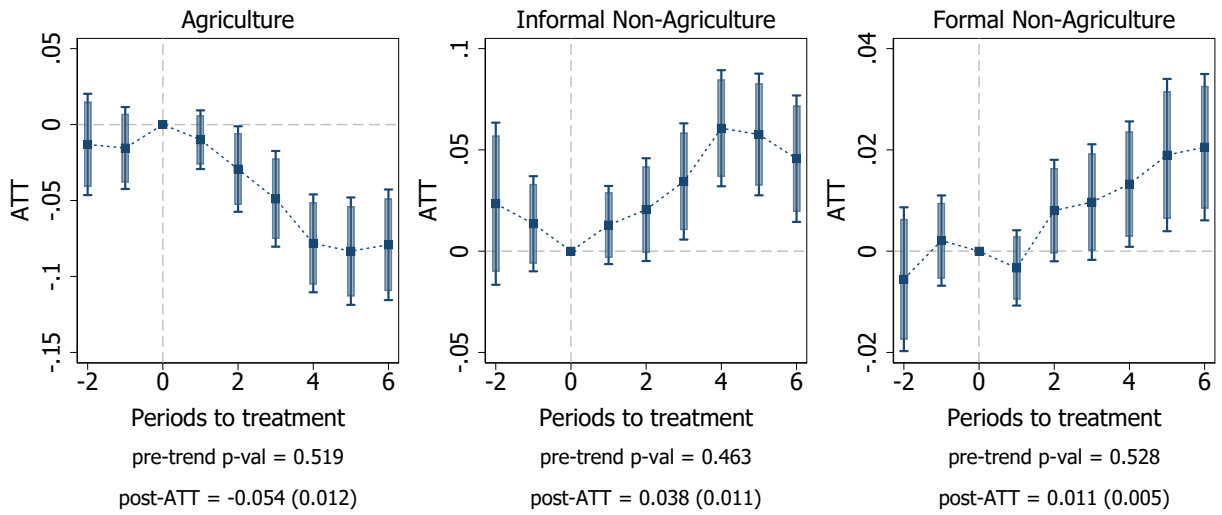
Sample includes households without high school graduates



Notes: This figure shows the effects of industrial zone exposure on the distance to a balanced labor portfolio, using data from VHLSS 2004–2020. The outcome is defined as distance-to-balance =  $\sum_j (l_j - 1/3)^2$  where  $l_j$  is the share of household labor hours in sector  $j \in \{A, F, I\}$ . Square markers indicate the point estimates of the coefficients. Darker vertical lines with caps show 95 percent confidence intervals, and lighter bars represent 90 percent confidence intervals. *pre-trend* p-val is the p-value from the joint test that pre-treatment effects are zero. *post-ATT* represents the average treatment effect on the treated across post-treatment periods, with standard errors clustered at the district level in parentheses. Estimates are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

Figure B4: Industrial Zones and Household Labor Share (Hours)

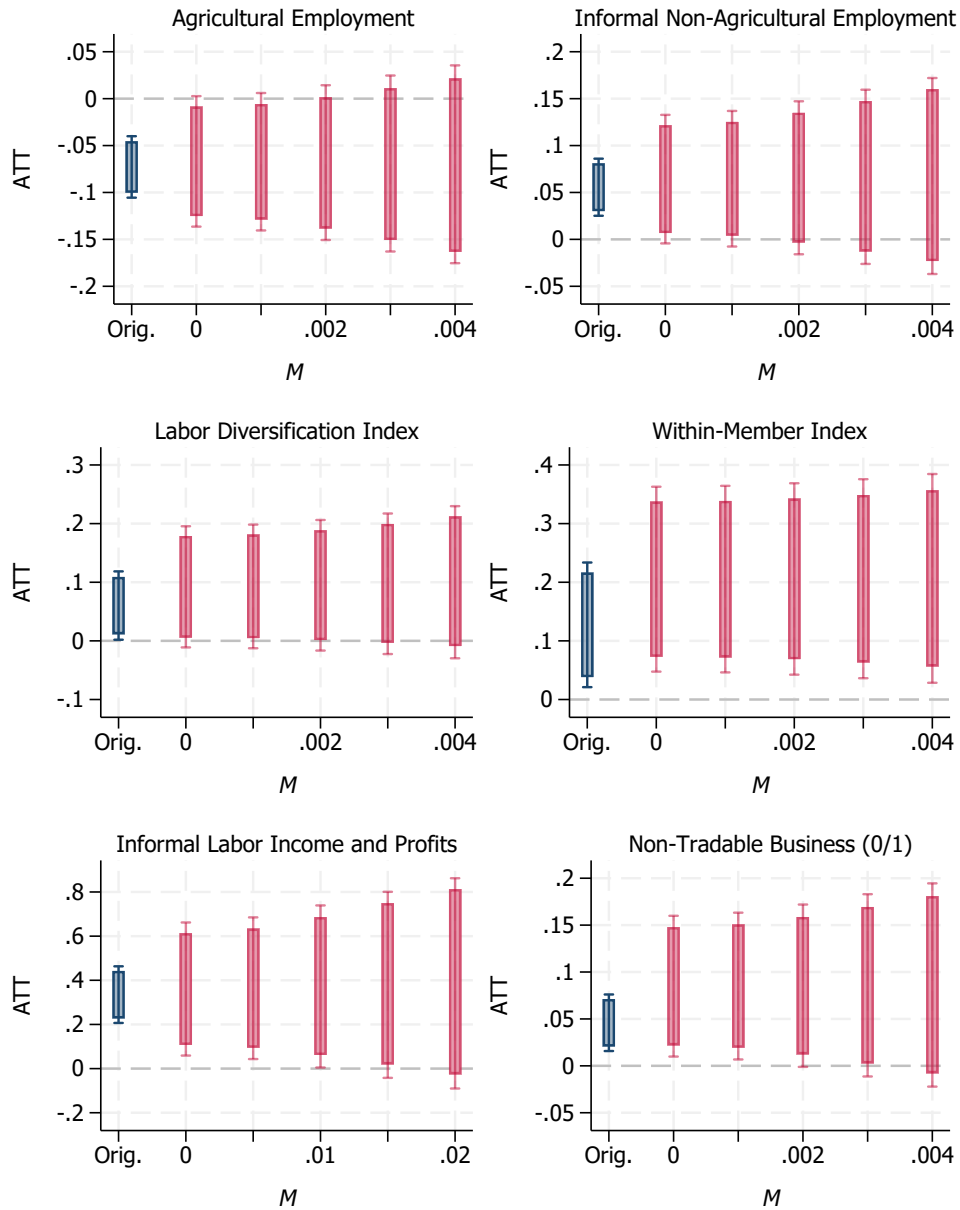
Sample includes households without high school graduates



Notes: This figure shows the effects of industrial zone exposure on the share of household total labor hours in each sector, using data from VHLSS 2004–2020. Square markers indicate the point estimates of the coefficients. Darker vertical lines with caps show 95% confidence intervals, and lighter bars represent 90% confidence intervals. *pre-trend* p-val is the p-value from the joint test that pre-treatment effects are zero. *post-ATT* represents the average treatment effect on the treated across post-treatment periods, with standard errors clustered at the district level in parentheses. Estimates are derived using the method proposed by de Chaisemartin and d’Haultfoeuille (2024). Sampling weights are applied throughout.

Figure B5: Industrial Zones and Selected Outcomes

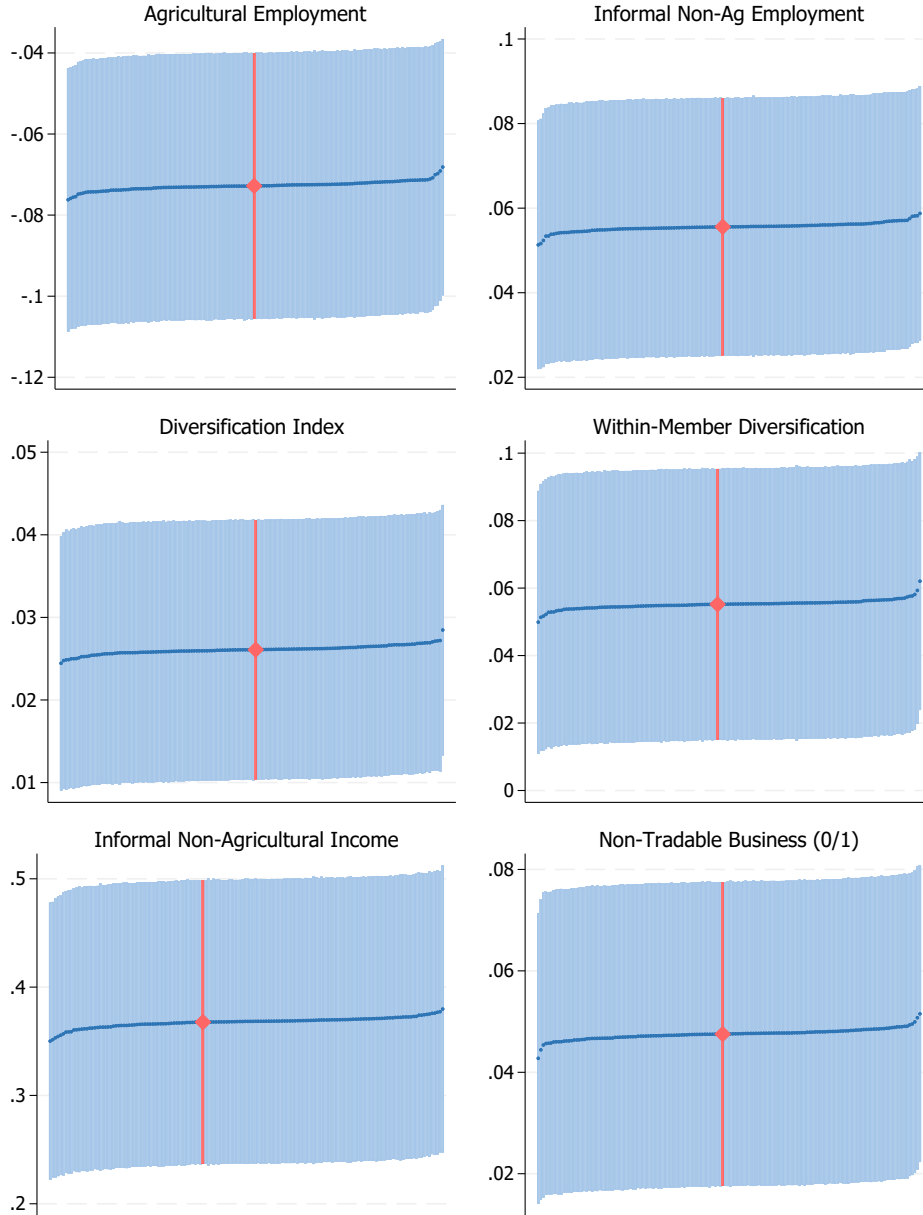
Robustness to Potential Violations of Parallel Trend Assumptions



Notes: This figure reports results from the sensitivity analysis proposed by Rambachan and Roth (2023), constructing robust confidence sets at period 4 post-treatment that allow for non-linear violations of the parallel trends assumption. Blue represents the original confidence intervals (90% and 95%) for the post-treatment average effect. The red represents fixed-length confidence intervals (FLCIs) for different values of  $M$ , where  $M = 0$  allows for linear violations of parallel trends (i.e., extrapolating a linear trend from the pre-treatment period), and larger values of  $M$  permit greater deviations from linearity. The sample includes individuals 25–64 without a high school diploma (employment outcomes), or households without any members completing high school (household diversification, income, and business outcomes).

Figure B6: Sensitivity Test: Industrial Zones and Selected Outcomes

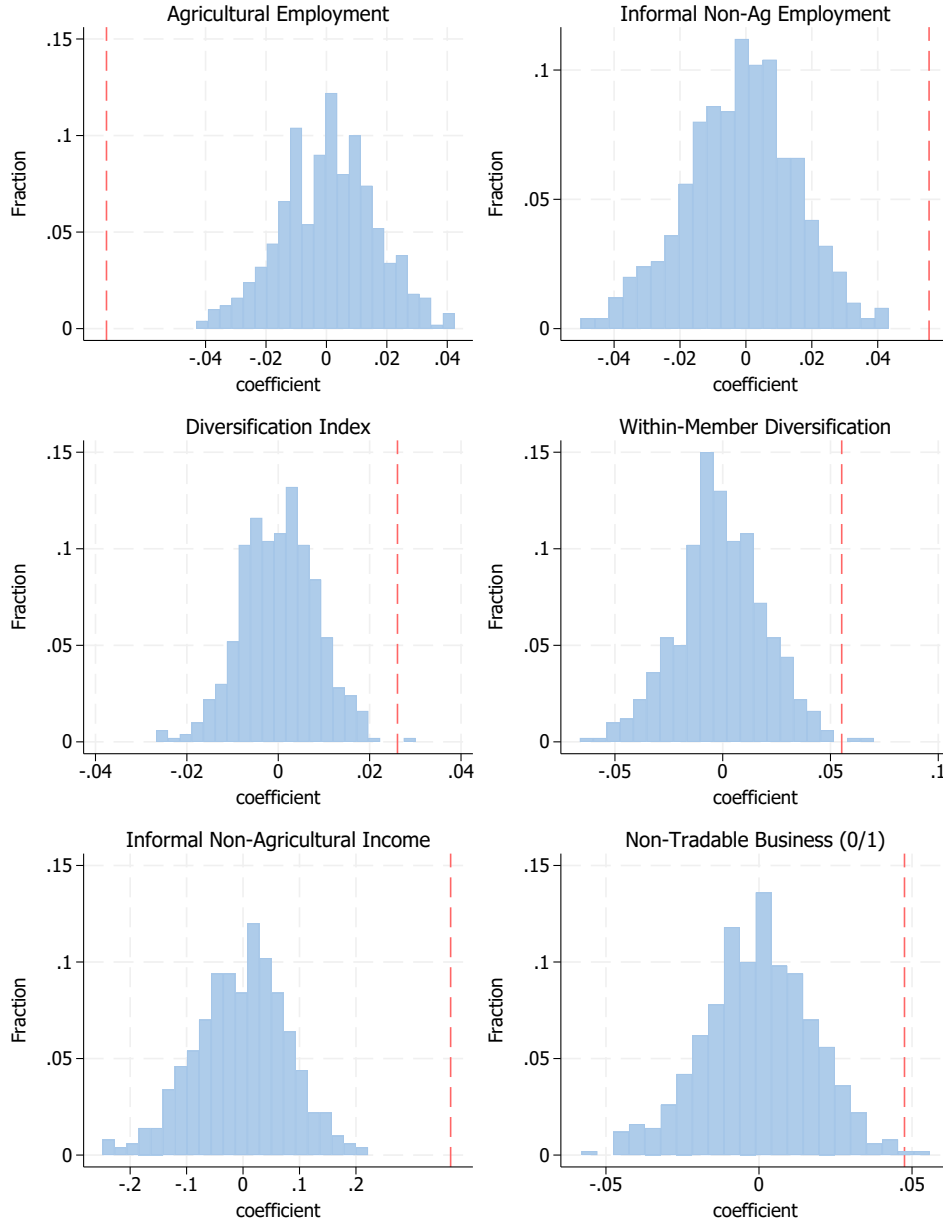
Treatment Effect of Period-4 Post-Zone Establishment  
 Results from Leave-One-District-Out Exercises



Notes: Each panel displays the distribution of estimated coefficients (blue circles) and their 95% confidence intervals from a leave-one-district-out exercise, in which each ever-treated district is sequentially excluded from the sample and the baseline specification re-estimated. The six outcomes are: agricultural employment and informal non-agricultural employment among working-age individuals (25–64) without a high school diploma (top row), household labor diversification index and within-member diversification among households without a high school graduate (middle row), and informal non-agricultural income and non-tradable business formation among households without a high school graduate (bottom row). Red diamonds indicate the baseline estimates and red vertical lines show the corresponding 95% confidence intervals. All estimates correspond to period 4 post-zone establishment. Sampling weights are applied throughout.

Figure B7: Placebo Test: Industrial Zones and Selected Outcomes

Treatment Effect of Period-4 Post-Zone Establishment  
Results from Monte-Carlo Permutation Tests



Notes: Each panel shows the distribution of estimated coefficients (blue bars) from 500 Monte Carlo simulations in which the treatment status time series of one district is randomly reassigned to another district's data and the baseline specification re-estimated. The six outcomes are: agricultural employment and informal non-agricultural employment among working-age individuals (25–64) without a high school diploma (top row), household labor diversification index and within-member diversification among households without a high school graduate (middle row), and informal non-agricultural income and non-tradable business formation among households without a high school graduate (bottom row). All estimates correspond to period 4 post-zone establishment. The red dashed vertical lines indicate the baseline estimates. The observed Type I error rates evaluated at 5% significance level range from 5.0% to 7.2% across outcomes, and randomization-based p-values are below 0.01 for all outcomes. Sampling weights are applied throughout.

## C DISTANCE-BASED ANALYSIS

In the main analysis, a district is classified as treated in the first year it falls within a 15-kilometer buffer of a newly established industrial zone. While the 15-kilometer cutoff is motivated by evidence from other contexts—for example, prior studies report statistically significant employment effects within a 10–15 km radius of special economic zones in India (Gallé et al., 2024), and wealth effects within a 10 km radius in several African countries (Abagna et al., 2025), it remains important to verify empirically that effects concentrate within this radius and do not spill over to more distant districts used as controls.

If spatial spillovers exist (outcomes in untreated districts are affected by their proximity to treated districts) then these “control” units are not valid counterfactuals, potentially violating the Stable Unit Treatment Value Assumption (SUTVA) and leading to biased treatment effect estimates. To assess the spatial reach of zone effects, I follow the existing literature (e.g., Gallé et al., 2024) and conduct a distance-based analysis. For each cohort of zone establishment  $c \in \{2004, 2006, 2008, 2010, 2012, 2014\}$  (representing more than 95% of zone establishment in the study period, Appendix Table B1), I restrict the sample to districts within that cohort and districts that were never treated under the 15-kilometer definition, and estimate the following specification:

$$y_{idt} = \gamma_d + \gamma_t + \sum_{b=1}^B \delta_b^c \cdot \text{POST}_t \times \text{Distance}_{d=b} + X'_d \phi \cdot t + \varepsilon_{idt} \quad (\text{C1})$$

where  $y_{idt}$  represents the outcome of interest for individual or household  $i$  in district  $d$  in year  $t$ . The terms  $\gamma_d$  and  $\gamma_t$  represent district and year fixed effects, respectively. The variable  $\text{POST}_t$  is a binary indicator equal to 1 for years after zone establishment in cohort  $c$ , and 0 otherwise. Because all treated districts within a cohort receive treatment at the same time, this avoids the negative weighting problem associated with staggered two-way fixed effects estimation (Borusyak et al., 2024; de Chaisemartin & d’Haultfoeuille, 2020; Goodman-Bacon, 2021). The variable  $\text{Distance}_{d=b}$  indicates whether district  $d$  falls within distance bin  $b$  from the nearest industrial zone:  $b = 1$  includes zone-hosting districts and those within 5 km,  $b = 2$  corresponds to 5–15 km,  $b = 3$  to 15–25 km,  $b = 4$  to 25–35 km, and  $b = 5$  to 35–45 km.<sup>5</sup> The omitted category is  $b = 6$  (45–70 km), which

---

<sup>5</sup>Because a district may fall within different distance bins for different industrial zones, I assign each district to the bin corresponding to the minimum distance to the nearest zone.

serves as the reference group. The vector  $\mathbf{X}_d$  includes baseline district characteristics (including share of urban population, share of ethnic minority, share of rice cultivated area, and share of individuals working in farming), interacted with a linear time trend to account for differential trends across districts with different initial conditions. Each coefficient  $\delta_b^c$  captures the effect of zone proximity at distance  $b$  relative to the most distant districts.

To obtain an overall estimate, I aggregate the cohort-specific coefficients using a weighted average:

$$\bar{\delta}_b = \frac{\sum_c n_c \cdot \hat{\delta}_b^c}{\sum_c n_c} \quad (\text{C2})$$

where  $n_c$  is the number of treated districts in cohort  $c$ , so that cohorts with more treated districts receive proportionally greater weight. Standard errors for the weighted average are constructed by combining the cohort-specific clustered standard errors. Since each cohort-specific regression uses non-overlapping sets of treated districts, I treat the cohort-specific estimates as approximately independent and compute the weighted average standard error as:

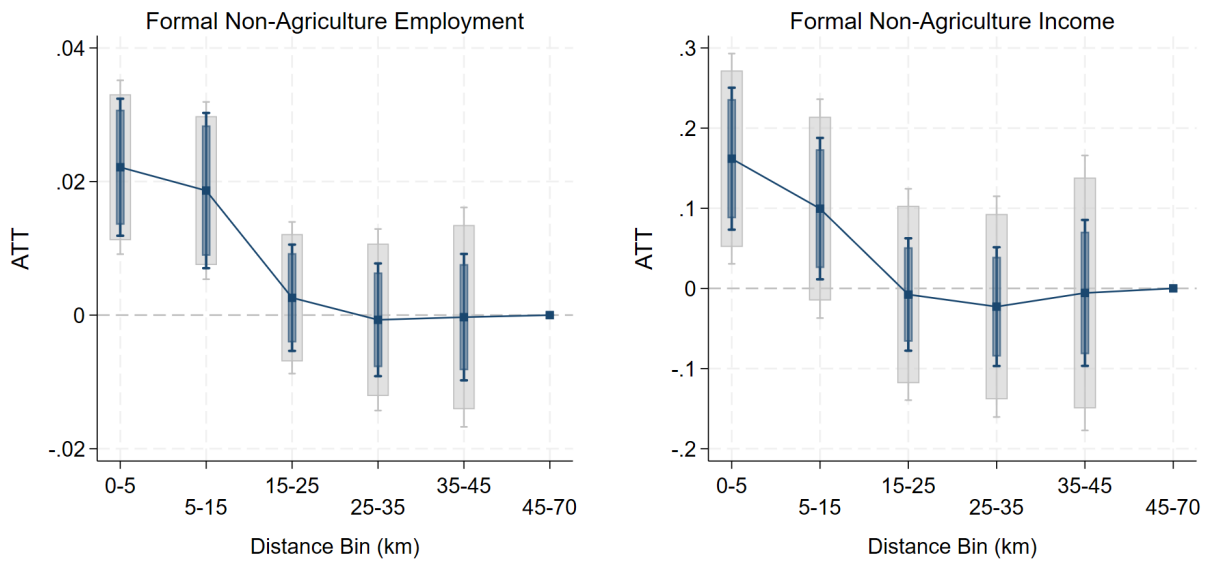
$$SE_{avg,b} = \frac{\sqrt{\sum_c n_c^2 \cdot SE_{c,b}^2}}{\sum_c n_c} \quad (\text{C3})$$

where  $SE_{c,b}$  is the clustered standard error for bin  $b$  from cohort  $c$ 's regression. Because all cohort-specific regressions share a common pool of never-treated control districts, the cohort-specific estimates are not strictly independent. This shared control group induces a positive covariance across cohort estimates, which the independence assumption ignores. The resulting standard errors may therefore be slightly understated. As a robustness check, I employ a stratified cluster bootstrap at the district level: in each of 200 replications, I draw districts with replacement separately within each treatment cohort and the never-treated group, preserving the cohort composition of the original sample, re-estimate equation (C1) for each cohort, and compute  $\bar{\delta}_b$  as in equation (C2). The standard deviation across replications serves as the bootstrap standard error. This approach preserves the cross-cohort dependence structure and accounts for within-district clustering.

If effects attenuate beyond 15 km, this supports the treatment definition used in the main analysis and suggests that SUTVA violations are unlikely to bias the main results. The results presented in Appendix Figure C1 indicate that the magnitude of the estimated effects is largest

for districts closest to the industrial zones and decreases progressively with increasing distance. Across outcomes including formal non-farm employment and total formal non-agricultural income, the coefficients are positive and generally statistically significant within the 0–15 km range. However, these effects quickly diminish with each successive distance bin and generally become statistically indistinguishable from zero beyond 15 km. These findings highlight the localized nature of the benefits associated with proximity to industrial zones, reinforcing the existing literature that documents how the economic and social spillovers of place-based policies tend to be highly concentrated around the immediate vicinity of the zones (Abagna et al., 2025; Ehrlich & Seidel, 2018; Gallé et al., 2024; Tafese et al., 2025).

Figure C1: Industrial Zones and Selected Outcomes: by Distance Bin



Notes: Dark shaded areas represent 95% and 90% confidence intervals based on analytical standard errors, computed as the weighted combination of cohort-specific clustered standard errors under an approximate independence assumption. Light shaded areas represent the corresponding 95% and 90% confidence intervals from a stratified cluster bootstrap (200 replications), which accounts for cross-cohort dependence induced by the shared pool of never-treated districts. Symbols represent the weighted average coefficients on  $POST_b$  for bin  $b$  from equation (C1), with cohort weights proportional to the number of treated districts. The outcome is the employment share in formal non-agriculture for workers aged 25–64 without a high school diploma (left) and total household income from formal non-agriculture (right). Sampling weights are applied throughout.

The 15-km cutoff may not apply uniformly across all outcomes and subgroups. For formal non-agricultural employment among workers aged 25–64 with at least a high school diploma, for example, point estimates in the 15–25km range remain positive, though statistically insignificant.

This pattern suggests that using the 15km radius as the treatment definition may yield conservative estimates for this subgroup, as districts experiencing modest positive spillovers are included in the comparison group. For less-educated workers, the primary focus of this paper, effects concentrate sharply within 15km, and no such spillover is evident. As a robustness check, I exclude districts in the 15–25km range from the comparison group in a donut specification, ensuring that any residual spillovers beyond the treatment boundary do not contaminate the counterfactual.

## Appendix References

- Abagna, M. A., Hornok, C., & Mulyukova, A. (2025). Place-based policies and household wealth in africa. *Journal of Development Economics*, 176, 103482.
- Balassa, B. (1964). The purchasing-power parity doctrine: A reappraisal. *Journal of Political Economy*, 72(6), 584–596.
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.
- Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230. <https://doi.org/10.1016/j.jeconom.2020.12.001>
- de Chaisemartin, C., Ciccina, D., D’Haultfoeuille, X., Knau, F., Malézieux, M., & Sow, D. (2024). *Event-study estimators and variance estimators computed by the did\_multiplegt\_dyn command* (Available at SSRN).
- de Chaisemartin, C., & d’Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- de Chaisemartin, C., & d’Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.
- Ehrlich, M. v., & Seidel, T. (2018). The persistent effects of place-based policy: Evidence from the west-german zonenrandgebiet. *American Economic Journal: Economic Policy*, 10(4), 344–374.
- Gallé, J., Overbeck, D., Riedel, N., & Seidel, T. (2024). Place-based policies, structural change and female labor: Evidence from india’s special economic zones. *Journal of Public Economics*, 240, 105259. <https://doi.org/10.1016/j.jpubeco.2024.105259>
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2), 373–377.
- Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Samuelson, P. A. (1964). Theoretical notes on trade problems. *Review of Economics and Statistics*, 145–154.

Tafese, T., Lay, J., & Tran, V. (2025). From fields to factories: Special economic zones, foreign direct investment, and labour markets in vietnam. *Journal of Development Economics*, 103467.