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

Trinh Pham*

This Version: January 1, 2026

Abstract

Do industrial zones generate broad-based welfare gains in high-informality settings, or do benefits concentrate among educated workers with formal access? Exploiting Vietnam's staggered rollout of industrial zones and national household surveys, difference-in-differences estimates show that formal employment and wage gains accrue disproportionately to educated workers. Less-educated households instead gain mainly through higher profits from informal nonfarm enterprises, concentrated in locally consumed services. Despite unequal formal access, total income and non-food consumption rise by similar magnitudes across education groups, alongside higher school enrollment and lower youth labor. Evaluations focused only on formal jobs can understate total benefits and misstate distributional incidence.

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 providing excellent 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

Targeted place-based industrial policies—including special economic zones (SEZs) and SEZ-type industrial zones—have become a central instrument for governments seeking to stimulate investment and growth in economically lagging regions. By 2019, more than 5,400 SEZs were operating worldwide (United Nations Conference on Trade and Development, 2019). These programs are typically justified by their potential to attract firms, raise productivity, and create jobs in high-productivity tradable sectors. In practice, however, their distributional consequences are far less clear in low- and middle-income countries where a large share of the workforce lacks the education and credentials required for formal manufacturing employment.

This paper asks: In high-informality settings, do industrial zones generate broad welfare gains, or do they primarily benefit a narrow segment of educated workers who can access formal jobs? The answer has first-order policy relevance. If zones disproportionately raise formal wages for educated workers, they may widen inequality in countries already facing large education gaps. At the same time, governments devote substantial fiscal and infrastructure resources to zone programs; assessing whether they reduce poverty and expand opportunity requires evidence not only on formal job creation, but also on who gains and through what channels.

A key takeaway of this paper is that evaluating zone programs solely through the lens of formal employment inside the zone may miss important general-equilibrium channels. In many developing countries, households combine agriculture, wage work, and self-employment, and informal non-farm businesses supply locally consumed (largely non-tradable) goods and services. When zones raise local incomes through formal-sector expansion, they can generate demand spillovers that increase the returns to informal activities. Whether these spillovers are large enough to produce broad-based welfare gains—particularly for less-educated households—is an empirical question, and existing evidence remains limited on both distributional incidence and mechanisms in high-informality contexts.

To guide the analysis, I develop a simple conceptual framework that clarifies how industrial zone expansion affects household labor allocation and income in a local economy. The framework highlights two features central to the setting: (i) a segmented labor market in which less-educated workers face barriers to formal non-farm employment, and (ii) an informal non-farm sector producing locally consumed non-tradable goods and services. Industrial zone establishment expands the scale and productivity of the formal non-farm sector, raising formal wages and increasing formal employment opportunities, though access for less-educated workers remains constrained. Higher formal earnings increase local demand for non-tradables, raising the equilibrium return to informal non-farm activities. 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. The framework yields testable predictions for sectoral reallocation, household labor diversification, informal enterprise returns, and welfare-related outcomes across education groups.

I empirically test these predictions using Vietnam's large-scale expansion of industrial zones over the last two decades. Vietnam provides an especially valuable setting: informality remains widespread, households maintain diverse livelihood strategies, and industrial zones have expanded rapidly since the early 2000s as part of an increasingly place-based approach to industrial development. I exploit variation in the timing and geography of zone roll-out across districts 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 10 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 2002–2020, the design identifies how localized industrialization affects labor market outcomes and household economic behavior.

The analysis yields four main findings that are broadly consistent with the model's predictions. First, industrial zones induce substantial sectoral reallocation, but access to formal employment remains highly unequal: following zone establishment, individuals with a high school diploma are nearly twice as likely as less-educated individuals to be employed in formal non-agricultural jobs. Less-educated workers instead reallocate primarily out of agriculture into informal non-agricultural activities. Second, households without a high school graduate exhibit a pronounced but non-monotonic increase in labor diversification: diversification rises in the years after zone establishment, peaks around four survey waves after treatment, and then gradually declines; households with educated members show no comparable diversification response. Third, despite limited formal job access, less-educated households experience significant income gains driven almost entirely by higher profits from informal household enterprises concentrated in locally consumed services (e.g., food preparation and accommodation). Fourth, although the underlying channels differ by education, total household income and non-food consumption increase by similar magnitudes across education groups; these gains are accompanied by higher school enrollment and lower adolescent child labor. Taken together, the results imply that industrial zones can generate broad-based welfare improvements not only through formal job creation, but also through local demand spillovers that expand informal economic opportunities and support inter-generational gains.

Beyond the specific context of Vietnam, these findings speak directly to the policy design and evaluation of place-based industrial programs. If a substantial share of welfare gains accrues through the informal non-tradable sector, then assessments that focus narrowly on "formal jobs created" may understate total benefits and mis-characterize distributional incidence. The attenuating effects with distance from the zone further suggests that the siting of zones and complementary investments in connectivity and local service-economy capacity are central for determining how many households fall within the effective catchment area of a zone and whether low-education households can share in the gains.

This paper contributes to three strands of the literature. First, it highlights and provide evidence consistent with a mechanism that is often under-emphasized in evaluations of place-based industrialization in high-informality settings. Much of the empirical SEZ literature evaluates zones using formal-sector outcomes such as employment and wages (e.g., Gallé et al., 2024; Y. Lu et al., 2019; Wang, 2013). A more recent strand has begun to measure household welfare impacts (e.g., Abagna et al., 2025), but evidence remains limited on the mechanisms through which less-educated households benefit when formal labor markets are segmented. I show that, in Vietnam, even when formal wage gains concentrate among educated workers, less-educated households experience welfare improvements of similar magnitude, with gains concentrated in informal non-farm household enterprise profits in locally consumed services. These findings are consistent with the view that, in high-informality settings, informal activities can be an important adjustment margin through which local demand spillovers transmit formal-sector growth into broader welfare gains, thus speaking to work on the developmental role of informality (Fields, 2011; La Porta & Shleifer, 2014; Maloney, 2004; Ulyssea, 2020). A direct policy implication is that evaluating zone programs solely by formal job creation can miss important channels of impact and mis-characterize who gains.

Second, the paper provides new evidence on how households adapt to localized industrial development through within-household labor reallocation and diversification across sectors. Classic dual-economy models emphasize a sharp transition from agriculture to industry as the hallmark of development (Herrendorf et al., 2014; Kuznets, 1973; Lewis, 1954), but recent evidence suggests structural change is often gradual and uneven across demographic groups (McMillan et al., 2014; Merotto et al., 2018). I show that adjustment to zone expansion differs sharply by education: more educated individuals disproportionately access formal occupations, while less-educated individuals expand into informal non-farm activities. Moreover, households without educated members temporarily increase cross-sector labor diversification, consistent with households combining activities when access to stable formal jobs is imperfect and local non-farm opportunities expand. These results contribute to a growing literature on household responses to place-based shocks (e.g., Zhao & Qu, 2024) and highlight policy-relevant margins—such as household labor portfolio choice and informal enterprise activity—that are typically outside the scope of worker-level formal employment metrics.

Third, the paper shows that localized industrialization increases school enrollment and reduces adolescent child labor, consistent with an income-effect channel in which improved household resources relax constraints on human capital investment, even among households with limited formal sector access. Prior studies emphasize that schooling decisions depend on perceived returns to education and opportunity costs, including the trade-off with child labor (e.g., Atkin, 2016; Basu & Van, 1998; Edmonds, 2005; Jensen, 2010), and show that place-based policies can have heterogeneous effects on educational attainment (F. Lu et al., 2023).[1] I find that industrial zones increase school enrollment and reduce adolescent child labor for both household groups, suggesting that short-run gains—potentially including gains associated with informal sector opportunities—can finance human capital accumulation. Moreover, these schooling responses are somewhat

[1]Evidence from China shows that technology-oriented SEZs tend to increase educational attainment, whereas export-led zones may reduce it (F. Lu et al., 2023).

stronger in more skill-intensive districts, a pattern that is consistent with zones raising the perceived value of educational credentials in places where formal skill demand is higher, though this heterogeneity should be interpreted descriptively. This pattern is policy-relevant because it points to a plausible pathway through which place-based industrial policies may relax inter-generational constraints on formal-sector access, strengthening the case for complementary education and skill investments in zone catchment areas (Glewwe & Muralidharan, 2016).

The remainder of the paper proceeds as follows. Section 2 describes the institutional background and labor market context. Section 3 introduces a simple conceptual framework that generates clear, testable predictions and informs the data construction and empirical strategy outlined in Section 4. Section 5 presents the main empirical results and discusses their implications in light of the model's predictions. Section 6 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.[2] At the national level, the Ministry of Planning and Investment oversees the planning, regulation, and promotion of industrial zones. It ensures new zones are integrated into national and provincial master plans and formally approved by the Prime Minister.

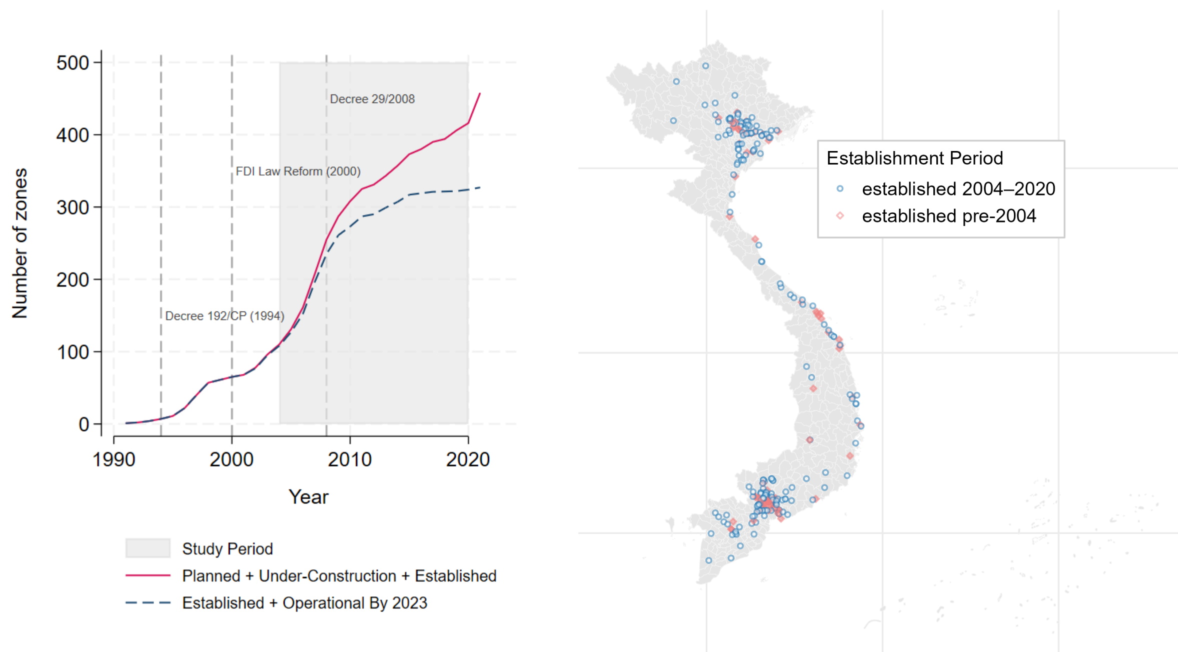
Provincial People's Committees play a central role in identifying potential locations, managing local implementation, and determining region-specific incentives and infrastructure investments. This decentralized yet collaborative model allows for top-down strategic direction while remain-

[2] 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 responsive to local development needs. Public and private investors can propose new zones, provided they meet the required infrastructure and regulatory standards. Once a zone is approved, provincial management boards, which operate under the Provincial People's Committees, are responsible for daily operations, administrative services, and monitoring compliance with investment and environmental regulations. This institutional structure helps ensure that industrial zones operate efficiently and remain aligned with national development priorities.

Figure 1 illustrates the growth trajectory of industrial zones in Vietnam from 1990 to 2020. The figure plots two key indicators: the cumulative number of planned, under-construction and established zones (solid red line) and cumulative zones that became operational by June 2023 (dashed blue line). The study period (2004–2020) is shaded in gray.

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 industrial zones in Vietnam by year. The red solid line indicates the total number of zones including planned, under-construction and established, while the blue dashed line represents cumulative zones that became operational by June 2023. The right panel displays the spatial distribution of industrial zones at both the beginning and the end of the study period. Data on individual zones are sourced from the Ministry of Planning and Investment.

The evolution of Vietnam's industrial zones reflects the influence of major policy reforms. By the mid-1990s, industrial zone development was limited, with fewer than 20 zones nationwide. This changed with the issuance of Decree 192/CP in late 1994, which 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, sustaining expansion at a more moderate pace. By 2020, over 400 industrial zones had been planned, with approximately 300 officially established and operational, reflecting Vietnam's ongoing strategy to strengthen its position as a leading industrial destination in the region.

The right panel of Figure 1 illustrates the spatial shift in industrial development. Zones established before 2002 (shown as red diamonds) 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 since 2002 (blue circles) are more geographically dispersed, extending into underdeveloped 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 industrial decentralization.

Table 1 documents substantial differences between firms located inside and outside industrial zones within the same district, using data from the Vietnam Enterprise Survey 2016. Industrial zones attract a markedly higher share of foreign-invested firms: nearly half of zone firms are foreign-owned, compared with only about 2% of firms outside zones. Firm size distributions also differ sharply. While firms outside zones are predominantly micro-enterprises with around 71% employ fewer than 10 workers, firms inside zones are substantially larger. More than 80% of zone firms employ at least 10 workers, and roughly 37% have 100 or more employees.

Industrial zones are overwhelmingly manufacturing-oriented. Manufacturing accounts for more than 80% of firms inside zones and nearly all zone employment, compared with roughly one-third of firms and employment outside zones. Despite this concentration, manufacturing firms inside zones do not appear to employ workers with systematically higher educational attainment than those outside zones. The share of manufacturing workers without a high school diploma is nearly identical across locations, reflecting the labor-intensive nature of zone-based manufacturing.[3]

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

	Outside Zones	Inside Zones	Difference (1) – (2)	
	Mean	Mean	Coef.	p-value
	(1)	(2)	(3)	(4)
Share of Domestic Firms	0.983	0.521	-0.462	0.000
Distribution by Firm Size				
Less than 10 employees	0.708	0.217	-0.491	0.000
10-50 employees	0.232	0.280	0.049	0.000
50-100 employees	0.030	0.143	0.113	0.000
100-300 employees	0.021	0.192	0.171	0.000
300 + employees	0.009	0.168	0.159	0.000
Distribution by Industry				
Agriculture	0.029	0.006	-0.023	0.000
Mining and quarrying	0.011	0.004	-0.007	0.000
Manufacturing	0.366	0.820	0.454	0.000
Public utilities	0.022	0.007	-0.015	0.000
Construction	0.172	0.027	-0.145	0.000
Sales, trade, hotels, restaurants	0.192	0.052	-0.140	0.000
Transports, storage, communication	0.076	0.031	-0.045	0.000
Finance, insurance, professional, business services	0.072	0.044	-0.028	0.000
Community, social, government services	0.060	0.008	-0.052	0.000
Share of Manufacturing Workers Without High School Diploma*	0.492	0.491	-0.001	0.977
Average Wage Bill of Manufacturing Workers (Million 2010 VND)	4.296	4.879	0.583	0.000

Notes: olumns (1) and (2) report mean values of the variables listed in the left-hand column for firms located outside and inside zones, respectively. Column (3) reports the coefficient on the inside-zone indicator from regressions of each left-hand-side variable 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 for the null hypothesis that the coefficient in column (3) equals zero. The industry distribution is weighted by firm size to reflect the distribution of employment across industries. The wage bill is defined as the firm's total payments to workers divided by the number of workers, with averages weighted by firm size. *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. Source: Calculations using the Vietnam Enterprise Survey 2016.

Wages, however, 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, implying a difference of about 14%. These are substantially higher than agricultural wages

[3]The Enterprise Survey does not directly record high school completion. 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. This classification is consistent with Vietnam's education system, in which these vocational programs do not require completion of upper secondary schooling.

of 2.8 million VND. This wage gap suggests that industrial zones may influence local economic conditions not only through employment concentration but also through higher earnings among manufacturing workers, with potential spillovers to surrounding areas.

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. In parallel, educational attainment increased markedly. School enrollment rose across all age groups, and the share of young adults attaining a high school diploma or higher increased substantially. By 2020, nearly all children aged 10–14 were enrolled in school, and the proportion of individuals aged 19–24 with a high school diploma rose by more than 20 percentage points relative to 2004.

Changes in the sectoral composition of employment differ sharply by educational attainment. Among workers without a high school diploma, the share employed in agriculture declined from 60% in 2004 to 34% in 2020, accompanied by rising employment shares in both informal non-agricultural activities (from 31% to 46%) and formal non-agricultural employment (from 8% to 20%).[4] Among workers with at least a high school education, agricultural employment also became less prevalent, declining from 23% to 9%, while formal non-agricultural employment accounted for an increasing share of employment, rising from 50% to 61%. In contrast, the share engaged in informal non-agricultural work remained relatively stable. Together, these patterns point to pronounced educational stratification in Vietnam's evolving sectoral employment structure, rather than direct movements of individuals across sectors.

This stratification becomes particularly relevant when considering potential spillovers from industrial zones. The expanding informal non-farm sector—comprised largely of elementary occupations in trade and sales, 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 14–18% reflect low barriers to entry and a high de-

[4] Agriculture primarily reflects informal agricultural employment, as formal agriculture accounts for a very small share of agricultural workers throughout the study period.

Table 2: Labor Market and Schooling Statistics, 2004–2020

Panel A: Labor Market Statistics

	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

Panel B: School Enrollment and Educational Attainment

	School Enrollment			Degree Attainment, 19-24	
	Age 10-14	Age 15–18	Age 19–24	High School	College/Uni
	(1)	(2)	(3)	(4)	(5)
2004	0.94	0.65	0.27	0.41	0.03
2012	0.96	0.69	0.35	0.57	0.08
2020	0.97	0.78	0.30	0.64	0.07

Notes: Panel A 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. Panel B reports school enrollment in the past 12 months by age group (Columns 1–3), and the share of individuals aged 19–24 who have attained a high school diploma (Column 4) or a post-secondary degree (Column 5). "Informality" is defined as employment in small-scale household businesses or self-employment. Sampling weights are applied throughout. Source: Calculations using VHLSS.

gree of fluidity across activities, particularly relative to self-employment in agriculture (McCaig & Pavcnik, 2021). These characteristics—accessibility to less-educated workers, low entry costs, and operational flexibility—make the informal non-farm sector 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 in settings where formal employment expands but remains unevenly accessible. 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. These mechanisms generate distinctive patterns of sectoral adjustment, household diversification, and welfare that guide the empirical tests. The full model, along with its assumptions and derived propositions and corollaries, is presented in Appendix A.

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:

$$l_A^s + l_F^s + l_I^s = 1, \qquad s \in \{H, L\} \tag{1}$$

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 θ that increases the scale and productivity of the formal tradable sector. In the simplest formulation, higher θ raises formal wages for both types, with a skill premium:

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

Here θ 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:

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

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

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

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:

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

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:

$$D_I(\theta, p_I) = Y_I(\theta, p_I) \tag{6}$$

This equilibrium condition is the central link between industrial zones and informal non-farm sector returns. When θ 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 θ. 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 services. 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 the local return 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.

(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.

Appendix A formalizes this framework, derives the comparative statics, and provides conditions under which these predictions hold.

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.[5] 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 650 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 biennially by Vietnam's General Statistics Office from 2002 to 2020. The survey employs a stratified multi-stage cluster sampling design, first selecting communes/wards—the third-level administrative units below provinces and districts—stratified by urban/rural status and geographic region,

[5] https://data.humdata.org/dataset/cod-ab-vnm

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 and labor compensations, and household measures of labor diversification, income and consumption expenditures.

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. Respondents report the industry, employer type, and hours worked for their two most time-consuming jobs. I restrict the sample to persons aged 25 to 64 to exclude most students and retirees.[6]

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.[7] 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 both across and within households. The first 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

[6]In 2002, the survey only collected information on the primary job. For most analyses that do not use the measure of skill intensity, which is constructed based on 2002 data, I begin with 2004 data.

[7]This enterprise-based definition (formal registration status) correlates highly (Pearson coefficient ≈ 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.

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 second 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.

The third 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 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.

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 labor and business 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.

Household Consumption Measures. The VHLSS collects detailed consumption data from a stratified sub-sample of approximately 9,000 households per wave. Food and non-food consumption modules capture both market purchases and self-produced goods. All food and non-food consumption values are expressed in per adult equivalent terms and deflated to 2010 Vietnamese Dong.[8]

Child Outcomes. For children aged 10-18, I examine both schooling and labor force participation. Similar to adults, a child is classified as working if they engaged in any income-generating activities in the past 12 months. School enrollment is measured using the Education Module, which records whether each child attended school during the same period. This allows me to analyze the trade-off between child labor and schooling. Additionally, education expenses (including tuition, textbooks, supplies, and tutoring) for each child, deflated to 2010 Vietnamese Dong, serves as a measure of household investment in human capital.

To identify the effects of industrial zones on these outcomes, 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 and treatment definition described in Section 4.2.

4.2 Empirical Strategy

Identifying the causal effects of industrial zones is challenging because zone placement is not exogenous. As shown in Section 2, zones established earlier were located in major economic centers with better infrastructure and more educated populations. Later zones spread to less developed regions, characterized by lower levels of urbanization, a higher proportion of ethnic minorities, greater reliance on agricultural employment, a smaller share of non-agricultural jobs, lower income from both wage work and household enterprises, and lower rate of children attending school (Appendix Table B1). These systematic differences in zone placement mean that simple comparisons between areas with and without zones would conflate the effects of zones with pre-existing characteristics or differential regional development trends.

[8] Adult equivalence is computed as number of AE = number of adults above 17 + 0.5 number of children 13 to 17 + 0.3 number of children 7 to 12 + 0.2 number of children 0 to 6 (De Janvry & Sadoulet, 2015).

Staggered DiD Design. To address this non-random placement challenge, I exploit the staggered rollout of zones across districts over time in a DiD framework and estimate an event-study specification as follows:

$$y_{idt} = \alpha_d + \delta_t + \sum_{k=-3,k\neq 0}^{6} \beta_k \cdot \mathbb{I}(t = \text{Establishment}_d + k) + \varepsilon_{idt} \tag{7}$$

where y_{idt} is the outcome of individual or household i in district d at time t. District fixed effects α_d absorb district-specific time-invariant characteristics, while year fixed effects δ_t control for nationwide year-specific shocks. The coefficients β_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 10 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 10-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 5-kilometer rings from zone centers shows that employment and welfare effects attenuate sharply beyond 10 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.[9]

Appendix Table B2 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 9.5% of districts establishing zones pre-2004, and 13.5% during 2004–2012. Zone creation then declined substantially, with only 1.4% during 2014–2020. The treatment exposure patterns follow a similar trajectory. Notably, 40% 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 β_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 β_k where $k \in \{-3, -2, -1\}$ 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.

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 in two ways. First, my spatial analysis shows that zone effects concentrate within

[9] 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.

the 10-kilometer buffer, with minimal impacts beyond this threshold. Second, robustness checks restricting the sample to long-term residents (using household registration data) yield similar results, suggesting that migration does not drive the findings (Appendix Section C).

Further Considerations. Three additional issues merit discussion. First, recent advances in the DiD literature has demonstrated that estimating the model using traditional two-way fixed effects (TWFE) can produce biased estimates when treatment effects vary across cohorts and overtime, as it implicitly uses early-treated units as controls for later-treated units and applies potentially negative weights to some comparisons (e.g., Borusyak et al., 2024; de Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021). This is particularly concerning in my setting for two reasons: (i) zones established at different times serve different economic purposes (United Nations Industrial Development Organization, 2019; World Bank Group, 2019), and (ii) early- and late-adopting districts differ substantially in baseline characteristics (Appendix Table B1).

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 a standard OLS. This approach compares districts with the same baseline treatment status to those that have not yet switched, ensuring that control groups remain untreated. It prevents contamination from already-treated units acting as controls and accommodates treatment effect heterogeneity across cohorts and over time. Placebo estimates for pre-treatment periods are used to assess the parallel trends and no-anticipation assumptions.

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 therefore run robustness checks by (i) re-estimating at the province level where representativeness is assured, (ii) re-estimating using household-panel data, and (iii) performing a leave-one-district-out analysis to test whether the findings are unduly influenced by any single district.

Finally, treatment and outcome timing do not perfectly align—zones can be established in any year, while household surveys occur biennially. I address this by assigning treatment based on the establishment year prior to each survey wave and testing robustness to alternative timing assumptions.

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.

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 among prime-age and older individuals (25–64), with post-treatment effects near zero (Appendix Figure B2).[10] While overall labor force participation remains stable, the sectoral composition of employment shifts significantly in response to industrial zone establishment.

Figure 2 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 post-treatment. Over the same horizon, the decline in agricultural employment is accompanied by increases in both formal and informal non-agricultural employment, with the expansion somewhat larger in the informal sector. These sectoral labor allocation responses are broadly consistent with the model's prediction.

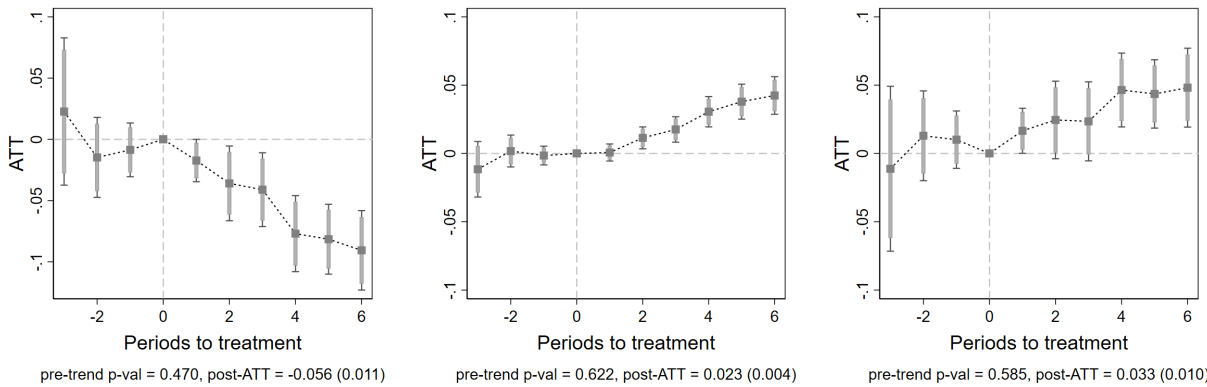
To better characterize the nature of these sectoral changes, Table 3 examines employment patterns across education groups using three complementary measures: sector of the primary job (Panel A), employment in a sector as either a primary or secondary job (Panel B), and employment in a sector as the sole job with no engagement in other activities (Panel C). Comparing these measures helps distinguish between complete exit from a sector, the addition of new activities alongside existing work, and the shedding of secondary activities—an important distinction in settings where multi-activity livelihoods are common (e.g., Barrett et al., 2001; Fields, 2011; Fox &

[10] The outcome is whether an individual engaged in any income-generating activity in the past 12 months, including agricultural or non-agricultural self-employment or wage employment. This null result indicates that behavioral changes documented below reflect transitions across sectors rather than changes in overall labor supply.

Figure 2: Industrial Zones and Primary Sector Employment

Workers Without High School Diploma

Agriculture | Formal Non-Ag | Informal Non-Ag



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 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.

Gaal, 2008).

Among individuals without a high school diploma, the results point to a pronounced reduction in agricultural employment across all three measures. Industrial zone exposure reduces agricultural employment by 5.6 percentage points when measured by the primary job (Panel A, Column 1; $p < 0.01$), corresponding to an approximately 9% decline relative to the pre-treatment mean of 63%. The magnitude remains similar when secondary jobs are included, with a 5.0 percentage point decline in Panel B ($p < 0.01$). Employment exclusively in agriculture—defined as having no engagement in other sectors—falls by 6.2 percentage points (Panel C, Column 1; $p < 0.01$). The consistency of these declines across panels suggests that the reduction reflects exits from agriculture rather than a reclassification of agriculture from primary to secondary activity.

At the same time, less-educated workers experience increases in both formal and informal non-agricultural employment. Formal non-agricultural employment rises by 1.5–2.4 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 3.3–4.1 percentage points ($p < 0.01$), corresponding to 10–15% increases relative to baseline shares of 19–31%. Notably, the presence of significant effects in both the "primary or secondary job" (Panel B) and "only job" (Panel C) measures indicates that informal employment expansion reflects entry as a primary source of livelihood rather than merely a supplemental activity alongside agricultural work.

Table 3: Industrial Zones and Sector Employment: Heterogeneity by Education Level

Sector	Without High School Diploma			With High School Diploma		
	Agriculture	Formal Non-Ag	Informal Non-Ag	Agriculture	Formal Non-Ag	Informal Non-Ag
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Primary Job						
Cumulative ATT	-0.056	0.023	0.033	-0.021	0.049	-0.028
	(0.011)	(0.004)	(0.010)	(0.016)	(0.017)	(0.015)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.63	0.05	0.31	0.24	0.48	0.28
p-value pre-trend	0.470	0.622	0.585	0.951	0.503	0.217
Panel B: Primary or Secondary Job						
Cumulative ATT	-0.050	0.024	0.041	-0.044	0.055	-0.026
	(0.012)	(0.004)	(0.010)	(0.019)	(0.016)	(0.016)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.77	0.07	0.40	0.41	0.49	0.35
p-value pre-trend	0.271	0.335	0.145	0.090	0.399	0.932
Panel C: Only Job (Single Sector)						
Cumulative ATT	-0.062	0.015	0.032	-0.020	0.040	-0.006
	(0.011)	(0.003)	(0.011)	(0.013)	(0.015)	(0.013)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.53	0.03	0.19	0.19	0.36	0.19
p-value pre-trend	0.095	0.210	0.440	0.954	0.223	0.223
N(Individual-Period)	270278	270278	270278	66353	66353	66353
N(District Switcher-Period)	880	880	880	874	874	874

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 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.

For individuals with at least a high school diploma, the adjustment pattern differs markedly. Prior to zone exposure, this group was already substantially less engaged in agriculture—24% in primary jobs, 41% in primary or secondary jobs, and 19% working exclusively in agriculture—compared with 63%, 77%, and 53%, respectively, among less-educated workers. When examining primary employment only (Panel A, Column 4), agricultural employment declines by roughly 2 percentage points, though the estimate is not statistically significant. When secondary agricultural activities are included (Panel B, Column 4), the decline becomes larger and statistically significant at 4.4 percentage points ($p < 0.05$). In contrast, employment exclusively in agriculture shows no significant change (Panel C, Column 4: -2 percentage points, $p > 0.10$).

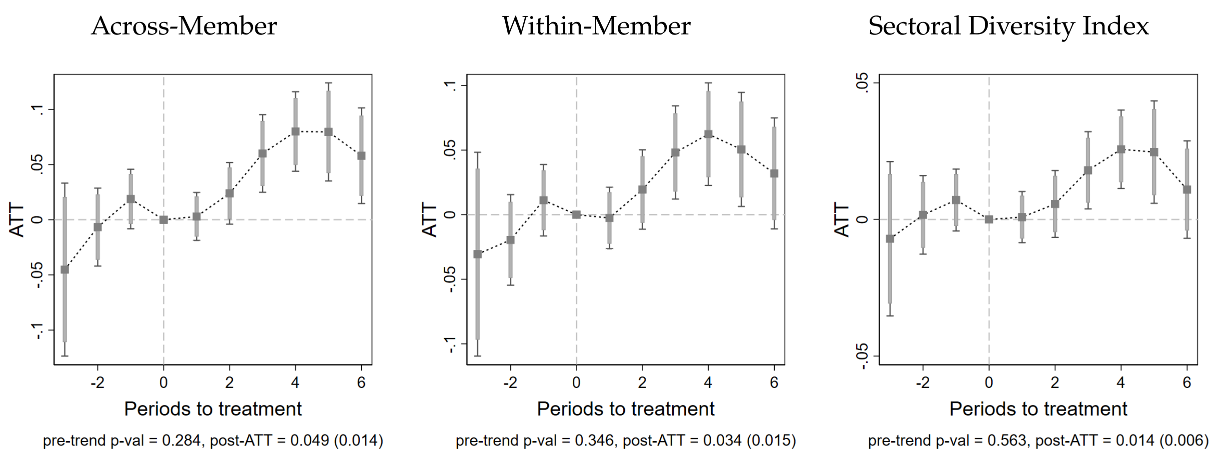
Taken together, these patterns indicate that more-educated workers reduce their engagement in agriculture primarily by discontinuing secondary or supplementary agricultural activities rather than exiting the sector entirely. Individuals specializing in agriculture appear to remain in the sector, while those combining agricultural and non-agricultural work drop the agricultural component of their livelihoods. Consistent with this interpretation, the main adjustment occurs within non-agricultural sectors. Formal non-agricultural employment increases by 4.0–5.5 percentage points ($p < 0.01$), corresponding to gains of roughly 10% relative to baseline levels. Informal non-agricultural employment, by contrast, declines modestly by 0.6–2.8 percentage points, with statistically significant effects only for primary jobs (Panel A; $p < 0.10$).

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—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. To assess whether household labor allocation responds along this margin, I next examine the effects of industrial zone exposure on labor diversification at the household level.

5.2 Household-Level Labor Diversification

Figure 3 plots event-study estimates of the impact of industrial zone exposure on three measures of household labor diversification for households with no high school graduate. Across all panels, the pre-treatment coefficients are small and the joint tests do not indicate differential pre-trends. After zone establishment, diversification rises gradually, becomes clearly positive within a few post-treatment periods, and peaks around periods 4–5 before attenuating. At the peak, across-member diversification increases by roughly 7–8 percentage points, within-member diversification by about 5–6 percentage points, and the sectoral diversity index by about 0.025. By period 6, the effects remain positive but are smaller and less precisely estimated (for the within-member and sectoral index measures, the confidence intervals include zero).[11]

Figure 3: Industrial Zones and Household Labor Diversification: Less-Educated Households



Notes: 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. Data from VHLSS 2004-2020. *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. Sampling weights are applied throughout.

To provide more insights into the diversification dynamics, Appendix Figure B3 Panel A plots event-study estimates for the distance of the household labor portfolio from an equal split of hours

[11] 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.

shares across agriculture, formal non-agriculture, and informal non-agriculture:

$$dist\_to\_balance = \sum_{j \in \{A,I,F\}} (l_j - 1/3)^2$$

Because $dist\_to\_balance$ 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, $dist\_to\_balance$ declines steadily and reaches its lowest point 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—$dist\_to\_balance$ 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.

Table 4 reports the cumulative effects on household labor diversification by education group. Households without any high school graduate experience economically meaningful and statistically significant increases in diversification across all three measures (Columns 1–3). Relative to pre-treatment means, these correspond to roughly 7–9% increases. In contrast, households with at least one high school graduate show no statistically detectable change in diversification: the estimates are small and imprecise across all metrics (Columns 4–6).

The stark divergence in household responses—less-educated households diversifying while more-educated households do not—is consistent with a long literature documenting 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 frictions in access to stable formal employment, leading households to combine activities across sectors when high-quality jobs are rationed (Harris & Todaro, 1970; Maloney, 2004; Ulyssea, 2020). The conceptual framework in Section 3 formalizes this latter mechanism: when formal opportunities expand but remain imperfectly accessible for less-educated workers, households adjust along

Table 4: Industrial Zones and Household Labor Diversification: Heterogeneity by Education Level

	Households Without High School Graduate			Households With High School Graduate		
	(1) Across Member Diversification	(2) Within Member Diversification	(3) Sectoral Diversity Index	(4) Across Member Diversification	(5) Within Member Diversification	(6) Sectoral Diversity Index
Cumulative ATT	0.049	0.034	0.014	0.015	-0.012	0.003
	(0.014)	(0.015)	(0.006)	(0.014)	(0.016)	(0.007)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.54	0.39	0.21	0.67	0.42	0.28
p-value pre-trend	0.284	0.346	0.563	0.511	0.171	0.814
N(Household-Period)	109932	109932	109932	67187	67187	67187
N(District Switcher-Period)	880	880	880	874	874	874

Notes: This table shows the effects of industrial zone exposure on household-level labor diversification strategies. In Columns (1) and (4), the dependent variable is whether the household has different members working in different sectors. In Columns (2) and (5), the dependent variable is whether the household has at least one member working in different sectors. In Columns (3) and (6), 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). 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. Sampling weights are applied throughout.

non-formal margins, reallocating labor toward informal activities whose returns rise with local demand.

Appendix Table B4 provides descriptive evidence on how these diversification effects vary with baseline local skill composition, measured using pre-treatment education shares in 2002. Across the three measures, diversification responses among low-educated households are larger in initially more skill-intensive districts. Consistent with this pattern, Appendix Figure B3 Panels B–C show that the portfolio distance to an equal split across agriculture, formal non-agriculture, and informal non-agriculture declines more in high-skill districts than in low-skill districts, indicating a stronger movement toward a balanced allocation of labor. However, interpreting this heterogeneity causally is challenging because baseline skill composition is correlated with multiple local characteristics, including initial formalization and urbanization, which could potentially affect the intensity and composition of zone-driven expansion. I therefore treat these patterns as descriptive heterogeneity—consistent with the model's broader implication that the incidence of zone impacts can vary systematically with local labor-market structure—rather than as a sharp test that isolates a single mechanism.

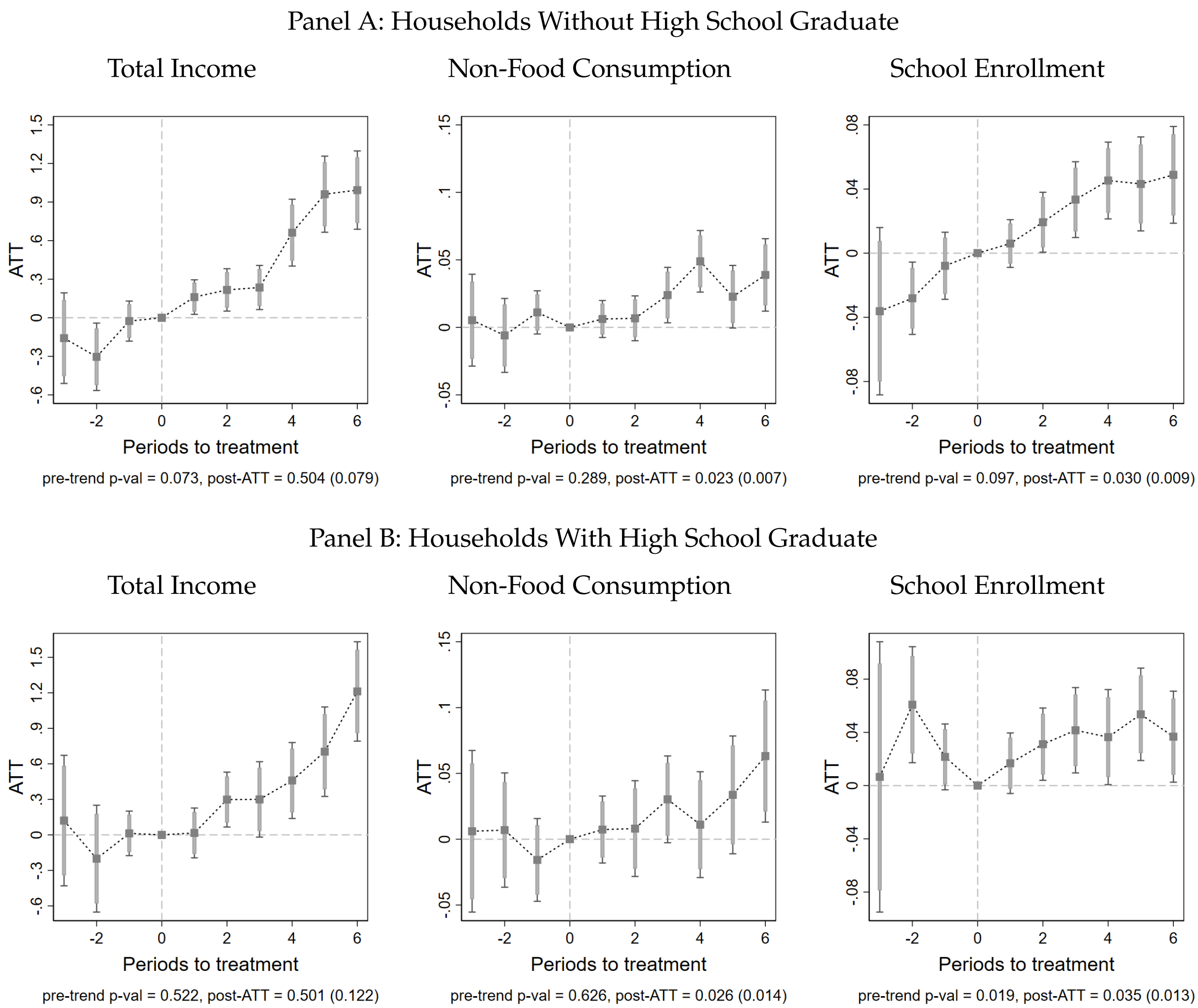
5.3 Economic Gains: Income, Consumption, and Human Capital Investment

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 economic gains, or do benefits concentrate among households able to access formal jobs? Guided by the conceptual framework, I next examine household income from wage work and household enterprises, and then consumption outcomes that capture welfare more directly. Because zone expansion can raise returns in locally consumed services and potentially local prices, I use consumption as well as downstream indicators of human capital investments to assess whether households experience real welfare improvements, and through which margins.

Figure 4 summarizes event-study estimates for three household living-standard indicators: total income, non-food consumption per adult equivalent, and school enrollment among adolescents aged 15–18. Despite the divergent labor-market adjustments documented above, both household groups exhibit positive post-treatment effects of similar magnitude. In particular, the post-treatment average effects are 0.504 (s.e. 0.079) for households without a high school graduate and 0.501 (s.e. 0.122) for households with a high school graduate for total income; 0.023 (0.007) and 0.026 (0.014) for non-food consumption; and 0.030 (0.009) and 0.035 (0.0.013) for enrollment of household members below 25 years old who have yet to complete high school education.

To understand how these gains arise, Table 5 decomposes household income by sector, separately by education group, using an unconditional specification that codes non-participation as zero (capturing both extensive- and intensive-margin adjustments). Two patterns stand out. First, for households without a high school graduate, income gains are concentrated in informal non-agricultural income, alongside more modest increases in formal non-agricultural earnings; agricultural income does not exhibit a statistically meaningful decline. Second, for households with a high school graduate, gains are driven primarily by formal non-agricultural earnings, with comparatively little change in informal income. Taken together, the decomposition is in line with the model's two-margin mechanism: formal-sector expansion raises formal earnings most directly for higher-education households, while less-educated households benefit disproportionately through increased returns to informal non-farm activities. Importantly, although the composition of gains differs sharply, the net increase in total income is similar in magnitude across education groups.

Figure 4: Industrial Zones and Household Welfare



Notes: 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. Household income is aggregated from wages earned by household members and profits from household enterprises across agriculture, informal and formal non-agriculture. Non-food consumption includes routine non-food household items and services, as well as education and health expenses per equivalent adult. Adult equivalence is computed as number of AE = number of adults above 17 + 0.5 number of children 13 to 17 + 0.3 number of children 7 to 12 + 0.2 number of children 0 to 6 (De Janvry & Sadoulet, 2015). School enrollment is for a sample of household members 24 years old and below that do not have a high school diploma. Data from VHLSS 2004-2020. *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. Sampling weights are applied throughout.

What accounts for the informal income gains among less-educated households? Disaggregating informal non-agricultural income shows that the increase is driven primarily by higher household non-farm business profits rather than informal wage earnings. Moreover, when I break down business performance by activity type, the largest and most precisely estimated gains are

Table 5: Industrial Zones and Sector Earnings: Heterogeneity by Education Level

Sector	Households Without High School Graduate				Households With High School Graduate			
	Agriculture	Formal Non-Ag	Informal Non-Ag	All	Agriculture	Formal Non-Ag	Informal Non-Ag	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative ATT	-0.098	0.171	0.408	0.504	-0.072	0.403	0.157	0.501
	(0.062)	(0.025)	(0.056)	(0.079)	(0.072)	(0.082)	(0.099)	(0.122)
District FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Mean Outcome	1.64	0.20	1.02	2.89	1.45	1.39	1.54	4.50
p-value pre-trend	0.172	0.472	0.312	0.073	0.545	0.077	0.170	0.522
N(Household-Period)	115030	115030	115030	115030	68222	68222	68222	68222
N(District Switcher-Period)	880	880	880	880	874	874	874	874

Notes: This table shows the effects of industrial zone exposure on household-level income from labor and business activities by sector and education level. The sample includes all households, coding non-participants as zero, capturing both extensive margin (entry into sector) and intensive margin (income changes). Income includes labor compensation (wages and benefits) earned by household members and profits from household enterprises. Agricultural income includes agricultural wages and household farm profits. Formal non-agricultural income includes wages from formal employment. Informal non-agricultural income includes informal wages and household non-farm enterprise profits. Income measures are winsorized at the 99th percentile within each survey year. All outcomes are measured in 2010 Vietnamese Dong. 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.

in locally consumed services: food and accommodation businesses experience sizable increases in profits (about 22%) and revenue (about 16%), whereas other informal activities show weaker or statistically insignificant changes (Appendix Table B5).

This pattern is consistent with the local-demand channel emphasized in the model: as industrial zones raise local earnings, demand for non-tradable services expands, increasing the returns to small-scale service activities that are accessible to less-educated households. At the same time, these results are reduced-form evidence and cannot fully rule out alternative channels such as differential entry/exit, changes in local competition, or measurement of business income across survey waves.[12]

Income gains translate into higher household consumption, with suggestive differences in composition by education group. Consistent with Engel-type patterns, less-educated households—who start from lower baseline consumption—allocate a larger share of additional resources to

[12]The absence of effects on informal manufacturing suggests limited supply chain spillovers to household enterprises. This may reflect the nature of informal household businesses in Vietnam: they typically employ fewer than 10 workers (often just the owner), operate with minimal capital, and lack quality certifications required by industrial zone firms. Evidence from SEZs in developing countries shows that formal firms primarily source key production inputs from imports or certified suppliers due to quality and scale requirements (e.g., Farole & Winkler, 2013; Javorcik, 2004).

food, whereas more-educated households show relatively larger increases in non-food categories such as education (Appendix Table B6). Notably, non-food consumption rises for both groups and the estimated magnitudes are similar, indicating that zone exposure is associated with broad improvements in material living standards even though the underlying income channels differ.

These gains also extend to human capital investment. For both household groups, industrial zone exposure increases school enrollment and reduces adolescent child labor among ages 15–18 as well as among youth 19–24 without a high school diploma (Appendix Table B7). A natural interpretation is an income effect: higher household resources relax budget constraints and reduce reliance on adolescents' labor (Basu & Van, 1998; Edmonds, 2005). A complementary (and not separately identified) interpretation is that zones may also raise perceived returns to schooling. Unlike settings where industrial expansion draws adolescents out of school into low-skill work (e.g., Atkin, 2016), the enrollment gains and child-labor declines in my setting are concentrated among youths near key educational transition ages and are stronger in more skill-intensive areas (Appendix Table B8)—patterns consistent with forward-looking schooling investment when industrial development increases the value of educational credentials (e.g., F. Lu et al., 2023). However, the relative importance of income versus returns channels cannot be isolated in this design.

5.4 Robustness Checks

In Appendix C, I conduct a comprehensive set of robustness checks to evaluate the stability of the main findings across three key outcome categories: (i) sectoral labor reallocation among working-age individuals, (ii) increased household labor diversification particularly households without a high school graduate, (iii) improved welfare outcomes across households.

For each category, results from alternative model specifications are reported to test various potential sources of bias and model sensitivity. These include: (1) varying the choice of comparison groups to include not-yet-treated in addition to never treated units; (2) using alternative inference approaches with different levels of error clustering to address spatial correlation; (3) assessing the Stable Unit Treatment Value Assumption (SUTVA) by investigating spillover effects based on district proximity to industrial zones; (4) controlling for demographic composition and migration effects, including restricting to individuals with permanent household registration to address sample composition concerns; (5) employing alternative staggered DiD estimators to accommodate differing parallel trend assumptions; (6) accounting for timing mismatches between annual treatment data and biennial outcome observations through split-sample estimation strategies; and (7) allowing for relaxation of parallel-trend assumption with district-specific linear time trends. Across all outcomes and alternative specifications, effects remain consistent in magnitude and statistical significance, reinforcing the credibility and robustness of the main results (Appendix Section C and Table C1).

As an additional check, I conduct a placebo test using Monte Carlo simulations of equation (7) to verify that the estimation approach yields unbiased and correctly inferred results. In each iteration, the entire district-level time path of industrial zone exposure is randomly reassigned across districts, so that one district's treatment history is applied to another district's individual- and household-level data, after which the treatment effects are re-estimated. Incorrect assignments should produce estimates with smaller magnitudes, zero means, or opposite signs. The baseline estimates lie outside the distribution of spurious estimates centered around zero, indicating that the observed effects are unlikely to occur by chance (Appendix Figure C2). The Type I error rates are approximately 4–6% at the 5% significance level, suggesting that the inference is fairly accurate against the null hypothesis of no industrial zone effect.

Next, I perform a leave-one-district-out analysis, sequentially excluding all observations from one district at a time and re-estimating the model. The results indicate that the estimated coefficients remain stable across iterations, with most estimates closely aligned with the baseline values, suggesting that the findings are not driven by any particular district (Appendix Figure C3).

Finally, a potential concern with the baseline analysis is that the VHLSS is designed to be representative at the province level, not the district level, so district-level estimates may be affected by compositional changes or sampling variation. While controlling for demographic characteristics and restricting the sample to long-term residents does not materially change the results (Appendix Table C1), I provide two additional evidence to address this concern.

First, I analyze the impact of industrial zones at the province level. Analyzing province-level effects presents two challenges. Most provinces have multiple industrial zones established over the study period—some with many zones predating the study—so the approach used in the district-level analysis is not appropriate. In addition, the smaller number of provinces (63 provinces versus

650 districts) raises concerns about statistical power. With these trade-offs in mind, I use the number of zones established as a continuous treatment variable and estimate an event-study specification following de Chaisemartin and d'Haultfoeuille (2024) and de Chaisemartin et al. (2024). The resulting estimates capture the average effect of one additional zone across all households in the province, including those far from industrial zones, and are thus expected to be smaller in magnitude than district-level estimates. As shown in Appendix Table C2, province-level results are qualitatively consistent with the baseline findings, supporting the robustness of the main results.

Second, given another advantage of the VHLSS being its rotating panel structure, i.e., the survey allows ones to track households for up to three consecutive survey waves (6 years), I construct a household panel by pooling four three-wave panels: 2004–2006–2008, 2010–2012–2014, 2012–2014–2016, and 2014–2016–2018. I then re-estimate the baseline specification in equation (7) on this sample to assess the robustness of the baseline results. The findings are qualitatively similar, indicating that changes in household demographic composition are unlikely to drive the main results (Appendix Table C2).

6 CONCLUSION

This paper studies how place-based industrialization affects labor-market adjustment and household welfare in a high-informality setting. Using the staggered rollout of industrial zones in Vietnam and a dynamic difference-in-differences design, I show that industrial zones reshape local economies through channels that extend beyond formal employment created inside the zone.

Guided by a simple conceptual framework with segmented access to formal jobs and a locally supplied non-tradable service sector, the empirical results highlight two key margins of adjustment. First, industrial zone exposure increases formal non-agricultural work disproportionately for more-educated individuals. Second, households with lower education levels reallocate labor out of agriculture and into informal non-farm activities, and their labor portfolios become more diversified in the years following zone establishment, with a non-monotonic (hump-shaped) profile over time.

These distinct labor-market pathways nonetheless translate into broadly similar improvements in key welfare proxies across education groups. Total income and non-food consumption rise for both less-educated and more-educated households, and adolescent outcomes improve: school enrollment increases and child labor falls among 15–18 year-olds. Decomposing income sources clarifies why similar aggregate gains can arise through different channels. For households without a high school graduate, income gains are driven primarily by higher profits from informal household enterprises, especially locally consumed services such as food and accommodation, whereas households with a high school graduate benefit mainly through higher formal non-agricultural earnings. This pattern is consistent with the framework's local-demand channel: formal earnings growth raises demand for non-tradable services, increasing returns to informal service activities that are accessible to less-educated households.

The findings have two policy implications for the design and evaluation of industrial zone programs in high-informality economies. First, assessments focused narrowly on formal jobs and wages may understate total benefits and mis-characterize distributional incidence, because a meaningful share of gains can accrue through indirect local spillovers that raise returns in the informal service economy. Second, complementary policies can shape how broadly zone-driven growth is shared: investments that ease constraints for small household enterprises (e.g., local infrastructure, market access, and business environment) may strengthen spillover gains, while education and skill investments and improved job matching can expand pathways into formal employment and reduce persistent barriers for less-educated workers.

More broadly, the paper underscores the importance of measuring multi-activity livelihoods and within-household labor reallocation when evaluating place-based industrial policies in the Global South. In contexts where households routinely combine wage work, self-employment, and agriculture, the welfare consequences of industrialization reflect both direct labor-market impacts and local general-equilibrium adjustments.

REFERENCES

Abagna, M. A., Hornok, C., & Mulyukova, A. (2025). Place-based policies and household wealth in africa. *Journal of Development Economics*, *176*, 103482.

Atkin, D. (2016). Endogenous skill acquisition and export manufacturing in mexico. *American Economic Review*, *106*(8), 2046–2085.

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.

Basu, K., & Van, P. H. (1998). The economics of child labor. *American Economic Review*, 412–427.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(6), 3253–3285.

de Chaisemartin, C., Ciccia, D., D'Haultfœuille, 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.

De Janvry, A., & Sadoulet, E. (2015). *Development economics: Theory and practice*. Routledge.

Dercon, S. (2002). Income risk, coping strategies, and safety nets. *World Bank Research Observer*, *17*(2), 141–166.

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

Edmonds, E. V. (2005). Does child labor decline with improving economic status? *Journal of Human rRsources*, *40*(1), 77–99.

Farole, T., & Winkler, D. E. (2013). *Making foreign direct investment work for sub-saharan africa: Local spillovers and competitiveness in global value chains*. World Bank Publications. http://documents.worldbank.org/curated/en/720931468203986757

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

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.

Glewwe, P., & Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the economics of education* (pp. 653–743, Vol. 5). Elsevier.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, *60*(1), 126–142.

Herrendorf, B., Rogerson, R., & Valentinyi, A. (2014). Growth and structural transformation. *Handbook of Economic Growth*, *2*, 855–941.

Javorcik, B. S. (2004). Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages. *American Economic Review*, *94*(3), 605–627.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, *125*(2), 515–548.

Kuznets, S. (1973). Modern economic growth: Findings and reflections. *American Economic Review*, *63*(3), 247–258.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, *28*(3), 109–126.

Lewis, W. A. (1954). Economic development with unlimited supplies of labour.

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

McMillan, M., Rodrik, D., & Verduzco-Gallo, Í. (2014). Globalization, structural change, and productivity growth, with an update on africa. *World Development*, *63*, 11–32.

Merotto, D. L., Weber, M., & Aterido, R. (2018). *Pathways to better jobs in ida countries: Findings from jobs diagnostics* (Vol. 1). World Bank Group. http://documents.worldbank.org/curated/en/675281538594680783

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.

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.

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.

Samuelson, P. A. (1964). Theoretical notes on trade problems. *Review of Economics and Statistics*, 145–154.

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.

United Nations Conference on Trade and Development. (2019). World investment report 2019: Special economic zones. 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

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 households 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:

$$\phi \coloneqq \frac{M_H}{M_H + M_L} \in (0, 1) \tag{A1}$$

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.[1]

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

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

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

[1] 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).

[2] Assuming a common returns-to-scale parameter α 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:

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

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:

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

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:

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

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 θ (e.g., more formal jobs available as the sector expands because of industrial zone establishment) but decreasing in ϕ (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 θ 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 θ 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 θ (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 θM 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. Note that this parameterization is consistent with the empirical finding that zone establishment increases formal employment by roughly 2.3 percentage points for less-skilled workers versus 4.9 percentage points for high-skilled workers (approximately a two-to-one differential effect, Table 3).[3]

[3] 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

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

so θ increases access by expanding vacancies while ϕ 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 θ 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 $l_j \to 0$, 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 θ, 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:

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

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

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

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

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

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

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

where the equality follows from Equation (A17). In other words, ω 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:

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

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

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

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:

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

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

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

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

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

And the total income is:

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

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

$$D_I = \frac{\mu Y^{total}}{p_I} \tag{A24}$$

Characteristics of Equilibrium. An equilibrium, given exogenous formal sector scale and productivity θ, 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 μ 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 $p_I \to 0$, nominal income from informal activities vanishes and total income converges to income from tradable sectors, which is strictly positive. Aggregate demand therefore diverges:

$$\lim_{p_I \to 0} D_I(p_I) = +\infty, \qquad \lim_{p_I \to 0} Y_I(p_I) = 0,$$

implying

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

As $p_I \to \infty$, 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:

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

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

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

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

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 θ. I examine how increases in θ influence:

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

[4] 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 θ 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:

$$\frac{dp_I}{d\theta} = -\frac{\partial \mathcal{H}/\partial \theta}{\partial \mathcal{H}/\partial p_I} \tag{A25}$$

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

- Supply effect: higher θ 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 θ 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 θ raises excess demand for informal non-farm goods and services at the initial price.

$$\frac{\partial \mathcal{H}}{\partial \theta} = \frac{\partial D_I}{\partial \theta} - \frac{\partial Y_I}{\partial \theta} > 0 \tag{A26}$$

Next, I consider the sign of $\partial \mathcal{H}/\partial p_I$. Holding θ 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:

$$\frac{\partial \mathcal{H}}{\partial p_I} = \frac{\partial D_I}{\partial p_I} - \frac{\partial Y_I}{\partial p_I} < 0 \tag{A27}$$

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 services, 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 θ leads to type-L households to reallocate non-formal labor away from agriculture and toward informal non-farm services.*

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

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

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

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*

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

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:

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

The total derivative is:

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

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

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

The sign of $\left[\bar{l} - (1-\bar{l})HHI_{A,I}\right]$ 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

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

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*

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

In particular,

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

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 ϕ gives

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

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 ϕ amplifies the θ-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 (θ, ϕ) 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 ϕ-amplification of the price-induced reallocation channel dominates the ϕ-attenuation of the direct formal-access channel, i.e.

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

where

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

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 ϕ 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 ϕ 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-ϕ 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 ϕ, 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 θ,

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

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

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

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):

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

By market clearing in the non-tradable sector,

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

so

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

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,

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

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

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

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: Characteristics of Districts in 2004

	Districts with Zones before 2004	Districts with Zones since 2004	Difference (2) – (1)	Districts with Zones 2004–2012	Districts with Zones 2013–2020	Difference (5) – (4)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Demographics						
Share of urban population	0.450	0.162	-0.288	0.159	0.162	0.002
	[0.433]	[0.228]	(0.044)	[0.221]	[0.206]	(0.049)
Share of ethnic minority population	0.026	0.091	0.065	0.111	0.074	-0.038
	[0.073]	[0.191]	(0.014)	[0.212]	[0.166]	(0.034)
Share of long-term registration	0.947	0.988	0.042	0.990	0.982	-0.008
	[0.066]	[0.016]	(0.007)	[0.013]	[0.021]	(0.006)
Panel B: Adult Labor Outcomes						
Share of agricultural labor	0.317	0.587	0.270	0.610	0.562	-0.048
	[0.261]	[0.192]	(0.028)	[0.189]	[0.205]	(0.051)
Share of formal non-agricultural labor	0.277	0.119	-0.158	0.111	0.129	0.018
	[0.184]	[0.091]	(0.019)	[0.087]	[0.078]	(0.019)
Share of informal non-agricultural labor	0.406	0.294	-0.112	0.280	0.310	0.030
	[0.143]	[0.143]	(0.016)	[0.146]	[0.163]	(0.043)
Income from wages and HH business profits	4.523	2.944	-1.579	3.009	2.743	-0.265
	[1.926]	[0.854]	(0.213)	[0.884]	[0.763]	(0.195)
Panel C: Household Labor						
HH labor diversification: across members	0.599	0.620	0.021	0.587	0.655	0.068
	[0.176]	[0.173]	(0.020)	[0.173]	[0.124]	(0.032)
HH labor diversification: within members	0.332	0.457	0.125	0.413	0.501	0.088
	[0.250]	[0.219]	(0.028)	[0.215]	[0.191]	(0.050)
HH labor diversification: HHI index	0.245	0.249	0.004	0.234	0.261	0.027
	[0.074]	[0.078]	(0.009)	[0.076]	[0.062]	(0.015)
Panel D: Child Outcomes						
Share of 10-18 attending school	0.840	0.801	-0.039	0.781	0.812	0.030
	[0.088]	[0.102]	(0.011)	[0.104]	[0.096]	(0.024)
Share of 10-18 working	0.194	0.283	0.088	0.290	0.289	-0.001
	[0.146]	[0.153]	(0.016)	[0.146]	[0.187]	(0.050)

Notes: This table summarizes 2004 characteristics across district types. Columns (1)–(3) compare early- versus later-treated districts: Column (1) includes districts within 10 km of a zone established before 2004; Column (2) includes those within 10 km of a zone established in 2004 or later; Column (3) reports mean differences. Columns (4)–(6) compare treatment timing within the study period: Column (4) includes districts within 10 km of a zone established 2004–2012; Column (5) includes those within 10 km of a zone established 2013–2020; Column (6) reports mean differences. Standard deviations in brackets; robust standard errors in parentheses. Sampling weights applied throughout.

Table B2: Establishment of Industrial Zones Across Districts

	Share of districts hosting zone (1)	Share of districts within 10-km radius of zone (2)
Pre-2004	0.095	0.322
2004-2012	0.135	0.245
2013-2020	0.014	0.029
Never-treated	0.756	0.404

Notes: This table presents the percentage of districts that either have an industrial zone within their boundaries (Column 1) or are located within a 10-kilometer radius of one (Column 2).

Table B3: Industrial Zones and Working Members

	Without High School Graduate (1)	With High School Graduate (2)
Cumulative ATT	-0.022	-0.023
	(0.035)	(0.049)
District FE	Y	Y
Year FE	Y	Y
Mean Outcome	2.81	2.81
p-value pre-trend	0.047	0.567
N(Household-Period)	115030	68222
N(District Switcher-Period)	880	874

Notes: The outcome is number of household members 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 Diversification: Heterogeneity by District-level Skill Intensity

Sample includes households without any members ever completing high school education

	Across-Member Diversification			Within-Member Diversification			Sectoral Diversity Index		
	All	Low	High	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative ATT	0.049	0.039	0.091	0.034	0.027	0.055	0.014	0.010	0.037
	(0.014)	(0.015)	(0.039)	(0.015)	(0.016)	(0.037)	(0.006)	(0.006)	(0.016)
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean Outcome	0.54	0.54	0.58	0.39	0.39	0.41	0.21	0.21	0.22
p-value pre-trend	0.284	0.660	0.348	0.346	0.477	0.082	0.563	0.841	0.223
N(Household-Period)	109932	93218	11396	109932	93218	11396	109932	93218	11396
N(District Switcher-Period)	880	667	183	880	667	183	880	667	183

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)–(3), the dependent variable is whether the household has different members working in different sectors. In Columns (4)–(6), the dependent variable is whether the household has at least one member working in multiple sectors. In Columns (7)–(9), 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). Columns (1), (4), and (7) include all districts; Columns (2), (5), and (8) include "low" districts only; and Columns (3), (6), and (9) include "high" districts only. A district is classified as "Low" if its share of working individuals with a high school diploma or a post-secondary degree is below the 75^{th} percentile of the district-level distribution in 2002, based on data from the 2002 VHLSS; districts above this threshold are classified as "High." 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 Household Non-Farm Business Performance

Heterogeneity by Activities

	Profits = Revenue - Cost			Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Local Food Production and Services	Other Activities	All	Local Food Production and Services	Other Activities
Cumulative ATT	1.406	3.364	0.168	0.157	7.307	-1.722
	(1.123)	(1.264)	(1.227)	(2.587)	(4.106)	(3.049)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	16.45	15.62	16.68	30.89	44.54	27.19
p-value pre-trend	0.305	0.228	0.129	0.305	0.304	0.267
N(Business-Period)	41277	12164	28680	41277	12164	28680
N(District Switcher-Period)	861	700	839	861	700	839

Notes: This table shows the effects of industrial zone exposure on household informal (non-registered) non-farm business performance by activities. "Local food production and services" comprises food and beverage production and services, as well as lodging services because in 2004–2006, the data do not distinguish between these two groups. However, from 2008 onward, informal household lodging businesses represent only a small share (about 5%) of this group, so it is reasonable to interpret the estimated effects as pertaining primarily to local food production and services only. "Other activities" include manufacturing, retail trade, construction, transport services, and other household enterprises. All values are expressed in 2010 million VND and winsorized at the top and bottom 2.5% for each industry group-year. 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. Sampling weights are applied throughout.

Table B6: Industrial Zones and Household Consumption: Heterogeneity by Education Level

	Without High School Graduate			With High School Graduate		
	(1)	(2)	(3)	(4)	(5)	(6)
Category	Food	Non-food	Education	Food	Non-food	Education
Cumulative ATT	0.043	0.023	0.005	0.016	0.026	0.019
	(0.012)	(0.007)	(0.002)	(0.021)	(0.014)	(0.006)
District FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Mean Outcome	0.43	0.21	0.04	0.52	0.36	0.08
p-value pre-trend	0.000	0.289	0.006	0.012	0.626	0.418
N(Household/Individual-Period)	22829	22829	40062	13392	13392	15532
N(District Switcher-Period)	869	869	856	800	800	784

Notes: This table shows the effects of industrial zone exposure on household-level food and non-food consumption per equivalent adult, as well as education expenses for household members younger than 25 years old without a high school diploma. Food consumption includes both regular daily food and special-occasion meals. Non-food consumption covers routine household items and services—such as fuel, personal care products, small goods, travel, and cultural expenses. Education expenditures encompass all school-related costs, including tuition fees, textbooks and reference materials, general school supplies, and private tutoring for subjects in the official curriculum for the sample of individuals younger than 25 years old without a high school diploma (unit of analysis is individual-year). All values are expressed in 2010 million VND and winsorized at the top 1% for each year. 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. Sampling weights are applied throughout.

Table B7: Industrial Zones and Children: Heterogeneity by Education Level

	Without High School Graduate			With High School Graduate		
	Work	School	Education Expenses	Work	School	Education Expenses
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 15-18 years old						
Cumulative ATT	-0.031	0.056	0.009	-0.032	0.043	0.058
	(0.018)	(0.017)	(0.005)	(0.021)	(0.016)	(0.019)
Mean Outcome	0.53	0.49	0.06	0.23	0.83	0.16
p-value pre-trend	0.604	0.028	0.477	0.778	0.096	0.230
N(Individual-Period)	35707	34924	6907	22320	22226	4021
N(District Switcher-Period)	869	869	669	860	860	568
Panel B: 10-14 years old						
Cumulative ATT	-0.005	-0.003	0.012	-0.016	0.002	0.007
	(0.013)	(0.007)	(0.005)	(0.012)	(0.005)	(0.008)
Mean Outcome	0.13	0.91	0.05	0.06	0.99	0.09
p-value pre-trend	0.537	0.361	0.528	0.090	0.251	0.281
N(Individual-Period)	51905	51308	10205	18070	18002	3150
N(District Switcher-Period)	877	877	753	817	817	473
Panel C: 19-24 years old, without high school diploma						
Cumulative ATT	-0.019	0.028	0.002	-0.037	0.027	0.011
	(0.009)	(0.006)	(0.004)	(0.020)	(0.015)	(0.007)
Mean Outcome	0.89	0.04	0.01	0.81	0.09	0.01
p-value pre-trend	0.466	0.014	0.555	0.715	0.569	0.607
N(Individual-Period)	31724	29944	6148	7792	7519	1124
N(District Switcher-Period)	816	813	575	678	672	193

Notes: This table shows the effects of industrial zone exposure on participation in economic activities (Columns 1 and 4), school enrollment (Columns 2 and 5), and education expenses (Columns 3 and 6). A child is classified as participating in economic activities if, within the past 12 months, they engaged in household farm or non-farm work, or in wage employment outside the household. Education expenses (monthly, 2010 million VND) encompass all school-related costs, including tuition fees, textbooks and reference materials, general school supplies, and private tutoring for subjects in the official curriculum. 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 B8: Industrial Zones and Human Capital: Heterogeneity by District-level Skill Intensity

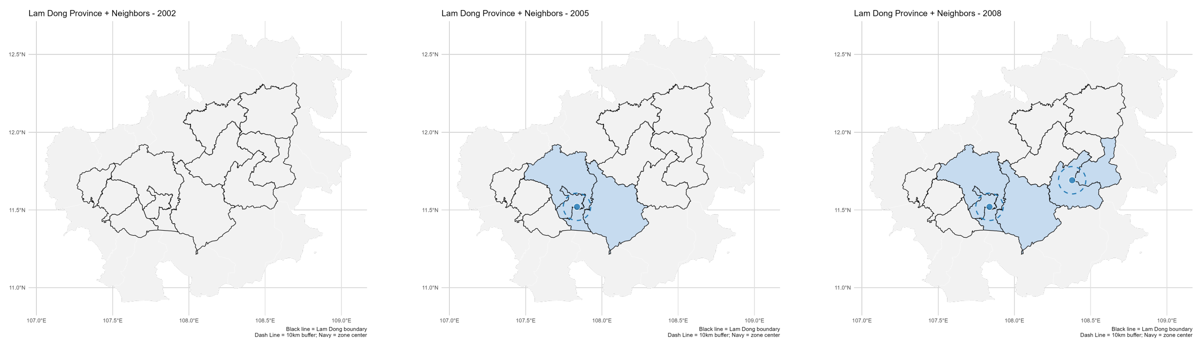
Sample includes individuals 15–18 and 19–24 without a high school diploma

	Work			School			Education Expenses		
	All	Low	High	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative ATT	-0.047	-0.039	-0.102	0.062	0.058	0.090	0.015	0.007	0.058
	(0.014)	(0.014)	(0.044)	(0.013)	(0.013)	(0.044)	(0.004)	(0.004)	(0.017)
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean Outcome	0.63	0.64	0.53	0.36	0.35	0.47	0.05	0.05	0.08
p-value pre-trend	0.375	0.088	0.722	0.413	0.138	0.980	0.159	0.042	0.032
N(Individual-Period)	91385	76488	9887	88445	74055	9729	18001	15082	2009
N(District Switcher-Period)	880	667	183	880	667	183	846	643	175

Notes: This table shows the effects of industrial zone exposure on labor participation (Columns 1–3), school attendance (Columns 4–6) and education expenses (Columns 7–9) across districts with different level of skilled labor demand. Columns (1), (4), and (7) include all districts; Columns (2), (5), and (8) include "low" districts only; and Columns (3), (6), and (9) include "high" districts only. A district is classified as "Low" if its share of working individuals with a high school diploma or a post-secondary degree is below the 75^{th} percentile of the district-level distribution in 2002, based on data from the 2002 VHLSS; districts above this threshold are classified as "High." 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.

B.2 Appendix Figures

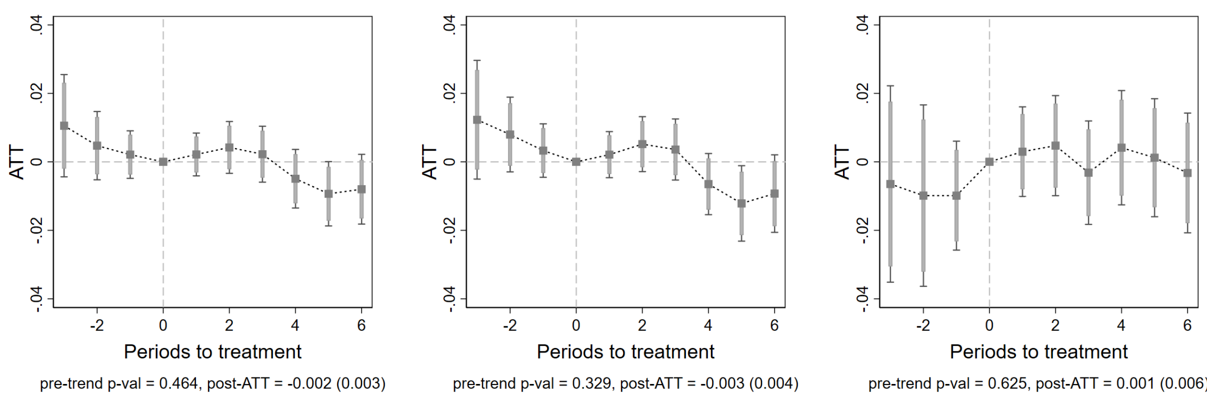
Figure B1: Mapping Zones and Identification of District Treatment Status



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 Labor Participation, Ages 25–64

(a) All (b) Without High School Diploma (c) With High School Diploma

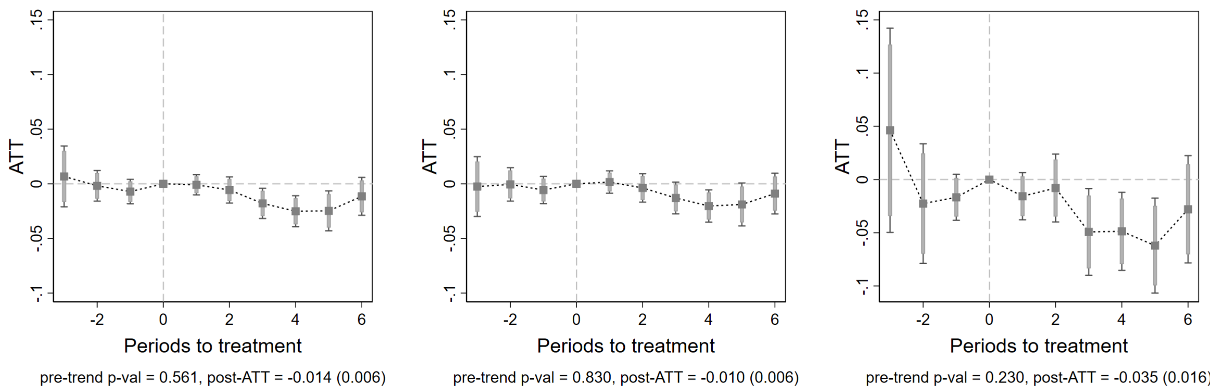


Notes: The sample includes individuals aged 25–64. Outcome is measured as whether an individual engaged in any economic activities during the last 12 months. The square symbol represents the point estimates of the coefficients. Darker vertical lines with caps indicate 95% confidence intervals, while the lighter bars represent 90% confidence intervals. Data from VHLSS 2004–2020. *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 B3: Industrial Zones and Distance to A Balanced Labor Portfolio

Households Without Any High School Graduates

Panel A: All Districts | Panel B: Low-Skill Districts | Panel C: High-Skill Districts

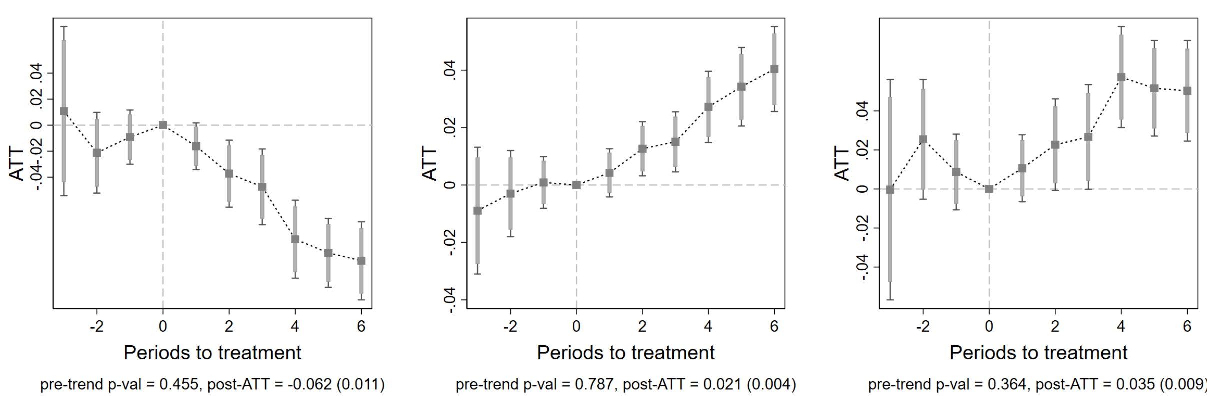


Notes: This figure shows the effects of industrial zone exposure on the distance to a balanced labor porfolio, using data from VHLSS 2004–2020. The outcome is defined as $dist\_to\_balance = \sum_j (l_j - 1/3)^2$ where l_j is the share of household labor hours in sector $j \in \{A, F, I\}$. Square markers 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)

Households Without Any High School Graduates

Panel A: Agriculture | Panel B: Formal Non-Ag | Panel C: Informal Non-Ag



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 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.

C ROBUSTNESS CHECKS

I conduct a comprehensive set of robustness checks to assess the stability of the main findings across three key outcome domains: (i) sectoral labor allocation among working-age individuals; (ii) increased household labor diversification; and (iii) improved household welfare outcomes, including income and children's schooling.

For each outcome domain, I report results from alternative model specifications designed to probe potential sources of bias and model sensitivity (Appendix Sections C.1–C.7). These checks include: varying the choice of comparison groups to incorporate not-yet-treated units in addition to never-treated units (Model 1); employing alternative inference strategies with different levels of error clustering to address spatial correlation (Model 2); controlling for demographic composition and migration effects, including restricting the sample to individuals with permanent household registration to address concerns about sample composition (Models 3 and 4); applying alternative staggered difference-in-differences estimators that allow for heterogeneous treatment timing and differing parallel-trends assumptions (Model 5); addressing timing mismatches between annual treatment data and biennial outcome measures through split-sample estimation strategies (Model 6); and relaxing the parallel-trends assumption by allowing for district-specific linear time trends (Model 7).

In addition, I implement a Monte Carlo permutation-based placebo test, in which the treatment time series of one district is reassigned to another district's employment and welfare outcomes, to verify that the estimation strategy does not mechanically generate spurious treatment effects under random assignment (Appendix Section C.8). I further conduct a leave-one-district-out analysis to examine whether the estimated effects are disproportionately driven by any single district (Appendix Section C.9). Finally, I complement the district-level analysis with province-level and household-level panel analyses to address potential concerns arising from district-level sampling variation and changes in household composition, given that the VHLSS is designed to be representative at the province level rather than the district level (Appendix Section C.10).

The results of these robustness checks are summarized in Appendix Tables C1–C2 and Appendix Figures C1–C3. In what follows, I discuss in details the motivation for each exercise and interpret the corresponding findings.

Table C1: Robustness Checks

	Baseline	Not-yet treated	Province clusters	Demo-graphic controls	Long-term residents	Callaway & Sant'Anna (2021)	Timing mismatch	District linear time trends
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A1: Outcome is working in agriculture as primary sector, 25–64								
Cumulative ATT	-0.044	-0.043	-0.044	-0.047	-0.045	-0.047	-0.053	-0.058
	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.021)	(0.043)
Panel A2: Outcome is working in agriculture as primary sector, 25–64, without high school diploma								
Cumulative ATT	-0.056	-0.054	-0.056	-0.057	-0.056	-0.055	-0.060	-0.082
	(0.011)	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)	(0.023)	(0.044)
Panel B1: Outcome is working in formal non-ag as primary sector, 25–64								
Cumulative ATT	0.026	0.025	0.026	0.026	0.025	0.020	0.041	0.030
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.021)
Panel B2: Outcome is working in formal non-ag as primary sector, 25–64, without high school diploma								
Cumulative ATT	0.023	0.022	0.023	0.022	0.022	0.017	0.036	0.019
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.015)
Panel C1: Outcome is working in informal non-ag as primary sector, 25–64								
Cumulative ATT	0.019	0.018	0.019	0.021	0.020	0.026	0.012	0.028
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.019)	(0.038)
Panel C2: Outcome is working in informal non-ag as primary sector, 25–64, without high school diploma								
Cumulative ATT	0.033	0.033	0.033	0.035	0.034	0.039	0.024	0.063
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.020)	(0.041)
Panel D: Outcome is across-member labor diversification, households without a high school graduate								
Cumulative ATT	0.049	0.049	0.049	0.047	0.048	0.037	0.049	0.118
	(0.014)	(0.014)	(0.016)	(0.014)	(0.014)	(0.013)	(0.014)	(0.060)
Panel E: Outcome is within-member labor diversification, households without a high school graduate								
Cumulative ATT	0.034	0.034	0.034	0.033	0.033	0.023	0.038	0.090
	(0.015)	(0.015)	(0.019)	(0.015)	(0.015)	(0.014)	(0.015)	(0.060)
Panel F: Outcome is sectoral diversity index, households without a high school graduate								
Cumulative ATT	0.014	0.014	0.014	0.013	0.014	0.011	0.014	0.043
	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.005)	(0.006)	(0.025)
Panel G: Outcome is total income from labor and business profits								
Cumulative ATT	0.571	0.541	0.571	0.562	0.578	0.387	0.571	0.485
	(0.093)	(0.093)	(0.093)	(0.091)	(0.093)	(0.079)	(0.093)	(0.354)
Panel H: Outcome is school enrollment of 15–24 without a high school diploma								
Cumulative ATT	0.062	0.057	0.062	0.037	0.063	0.053	0.058	0.059
	(0.013)	(0.013)	(0.014)	(0.010)	(0.013)	(0.012)	(0.022)	(0.043)

Notes: "Baseline" represents the estimate of equation (7), where the comparison group includes never treated units only, standard errors are clustered at the district level, using estimator by de Chaisemartin et al. (2024). Column (2) also includes not-yet-treated districts as comparison group. Column (3) is the same as baseline specification but clusters standard errors at the province level. Column (4) is similar to baseline specification but also controls for demographic characteristics including age, gender, and ethnic minority. Column (5) is similar to baseline specification but restricts the sample to long-term residents only. Column (6) employs a different staggered DiD estimator by Callaway and Sant'Anna (2021). Column (7) addresses concern with timing mismatches between annual treatment data and biennial outcome observations through split-sample estimation strategies. Column (8) allows for relaxation of parallel-trend assumption with district-specific linear time trends. Sampling weights are applied throughout.

C.1 The Choice of Comparison Group

The first robustness check assesses sensitivity to the choice of comparison group in the staggered DiD design. The baseline specification uses only never-treated districts as controls. This approach has the advantage of minimizing potential bias from compositional changes in the comparison group over time, in which not-yet-treated units eventually become treated, as well as the endogeneity that arises if districts' future treatment timing correlates with pre-treatment outcomes (Baker et al., 2025). However, the method raises the concern that never-treated districts may systematically differ from treated districts in unobserved ways (e.g., remaining untreated due to underlying outcome trends). As an alternative, I extend the comparison group to include not only never-treated districts, but also not-yet-treated districts. Reassuringly, results remain quantitatively similar to the baseline across all outcomes in both effect size and statistical significance (Appendix Table C1 Column 2).

C.2 Alternative Inferences

Second, the conclusions hold under alternative inference strategies. In the baseline specification, standard errors are clustered at the district level, where the treatment is assigned, to account for potential intra-district correlation. As a robustness check, I also cluster standard errors at the province level, the first administrative tier above the district and responsible for planning, managing, and coordinating industrial zone development. This adjustment captures potential spatial correlation in errors across districts within the same province. The consistency of the results across both clustering levels in Columns (1) and (3) of Appendix Table C1 increases confidence that the estimated effects are not driven by unobserved regional heterogeneity and are not sensitive to the clustering choice.

C.3 Demographic Composition and Migration Concerns

To capture the total impact of industrial zone exposure, including any migration induced by the zones, I estimate the baseline specification without controlling for any demographic variables. Although including these variables could improve precision, they risk serving as "bad controls" if they are themselves influenced by the treatment. In particular, if industrial zones attract migrants and thereby alter local demographics in ways related to the outcomes of interest, controlling for these variables could bias the estimated effects.

As shown in Appendix Table C1 Column (4), including controls for individual characteristics such as age, gender, and ethnic minority does not meaningfully change the estimated treatment effects. Across panels, the coefficients remain closely aligned with the baseline estimates, reinforcing the robustness of the findings.

To further address concerns about compositional changes, I restrict the analysis to individuals/households with long-term permanent registration status (*hộ khẩu thường trú*) in the same commune—a sub-district administrative level. During the study period, Vietnam operated a household registration system known as *hộ khẩu*, which determines individuals' official place of residence and conditions access to public services (Demombynes & Vu, 2016). Under this system, permanent registration signifies long-term residence and provides priority access to public services, including education, healthcare, and social assistance, within a locality; those with only temporary registration face significant barriers to these services. Because internal migrants often lack permanent registration in their destination areas, focusing on individuals with *hộ khẩu thường trú* could help isolate the treatment effect from short-term or unregistered migration, thereby mitigating potential composition bias.[5]

Restricting the sample to individuals who reside in the same dwelling as their registered permanent residence yields estimates that are similar in magnitude and statistical significance level to the baseline results across outcomes (Appendix Table C1 Column 5). These findings are perhaps not surprising given the earlier discussion of industrial zone rollout patterns in the country (Figure 1 and Table B1). Districts with zones established prior to the study period, typically located in major urban centers, tend to attract more internal migrants. On average, 5.3% of the population in these early-establishing districts lacks long-term household registration, though this figure varies considerably across provinces. For instance, in Ho Chi Minh City, 14% of individuals surveyed in 2004 reported not having household registration in the same commune; by 2020, that share had doubled to 28%. In Binh Duong, a major hub of industrial activity, the corresponding figures were 4% in 2004 and 55% in 2020.

In contrast, districts where industrial zones were established after 2004—the primary focus

[5] In the surveys, individuals are asked about their place of household registration: whether it is in the same dwelling within the commune/ward, elsewhere in the province, in another province, or if they have never registered.

of this study—are generally less developed and have lower rates of non-registered residents: just 1.2% in 2004, increasing modestly to 2.3% by 2020. These patterns suggest that most individuals in the analytic sample, especially in later-establishing districts, are long-term residents. This likely explains why limiting the sample to those with permanent residency has little effect on the estimated treatment impacts.

C.4 Alternative Staggered DiD Estimator

As another robustness check, I re-estimate the results using the estimator from Callaway and Sant'Anna (2021). The Callaway and Sant'Anna (2021) (CS) and de Chaisemartin and d'Haultfoeuille (2024) (dCDH) estimators diverge fundamentally in their parallel trends assumptions and control-group construction. The CS estimator relies on unconditional parallel trends for untreated potential outcomes, comparing treated units to never-treated units without conditioning on baseline treatment status and imposing parallel trends in post-treatment periods only. In contrast, dCDH requires conditional parallel trends within baseline-treatment cohorts, restricting comparisons to units with identical baseline treatment and unchanged status ("stayers"). Consequently, CS captures broad treatment effects but may be biased under baseline-heterogeneous trends, while dCDH provides cohort-specific estimates resilient to such heterogeneity at the cost of reduced generalizability.

The results in Appendix Table C1 Columns (1) and (6) show that the coefficient estimates and significance levels are largely consistent across both estimators for all outcomes, reinforcing the robustness of the main findings to alternative estimation strategies within a staggered adoption framework.

C.5 Mismatch Between Treatment and Outcome

A complication in the estimation arises from the mismatch in temporal frequency between treatment and outcome data: industrial zones are established annually, whereas VHLSS outcomes are observed biennially. The baseline strategy retrospectively aligns treatment status with survey years. For example, if a zone is established in 2009, the corresponding group is coded as treated in 2010, the next year in which outcomes are observed. This approach avoids using post-treatment outcomes to define pre-treatment periods and preserves a consistent structure. Under this framework, period 1 corresponds to 2 years after treatment, period 2 to 4 years after treatment, and so on. Furthermore, I focus on the post-treatment periods up to 12 years after treatment (period 6). This decision is driven by the availability of sufficient treated units across these periods. Starting from period 7, the number of treated units experiences a sharp decline, which could lead to imprecision in estimating treatment effects. Limiting the analysis to periods 1 through 6 ensures that the estimates are both precise and reliable.[6]

To validate the robustness of this timing assumption, I follow de Chaisemartin and d'Haultfoeuille (2024) and de Chaisemartin et al. (2024) and adopt a split-sample estimation strategy that leverages all available variation in treatment timing while maintaining the biannual frequency of the outcome data. Specifically, I estimate a subsample that includes all group-year observations in which the group's treatment status has remained unchanged since the beginning of the panel, as well as observations where the group experienced a treatment change in a year when outcomes are observed. Within this sub-sample, the event-study coefficients capture the effect of treatment exposure over odd-numbered intervals: the first coefficient reflects one biennial period of exposure, the second reflects three periods, and so forth. Column (7) of Appendix Table C1 show that this sub-sample generally yields similar estimates across outcomes, suggesting that the timing approximation in the baseline does not materially bias the estimated impacts.

C.6 District-specific Linear Time Trends

As another robustness check, I augment my baseline specification with district-specific linear time trends. This approach relaxes the parallel trends assumption by allowing treated and never-treated districts to follow differential trajectories. Results in Appendix Table C1 Column (8) show that point estimates for low-educated subgroups generally increase when district-specific trends are included, while the estimates for household income and child schooling are virtually identical to the baseline, and qualitative conclusions are robust throughout. This pattern suggests heterogeneous selection into treatment: districts establishing industrial parks were on less favorable trajectories specifically for low-skilled populations, while aggregate conditions remained relatively comparable across treated and control districts. For low-educated subgroups, the trend-adjusted

[6] This restriction does not change the direction of the effects. If anything, the magnitude of the average effects is somewhat larger when using the longer period restriction, as more periods are available to capture the full dynamic impact of the treatment.

estimates may therefore more accurately capture the policy's impact, as baseline estimates that assume parallel trends in levels could understate effects by attributing post-treatment gains partly to convergence toward faster-growing comparison districts. However, estimating district-specific trends requires many additional parameters, reducing precision across all estimates, and with only three pre-treatment periods, the counterfactual trajectories rest on limited pre-treatment information. I therefore treat the baseline as a lower bound for effects on low-skilled populations.

C.7 Distance-Based Analysis

In the main analysis, a district is classified as treated in the first year it falls within a 10-kilometer buffer of a newly established industrial zone. While the 10-kilometer cutoff may seem arbitrary, it 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).

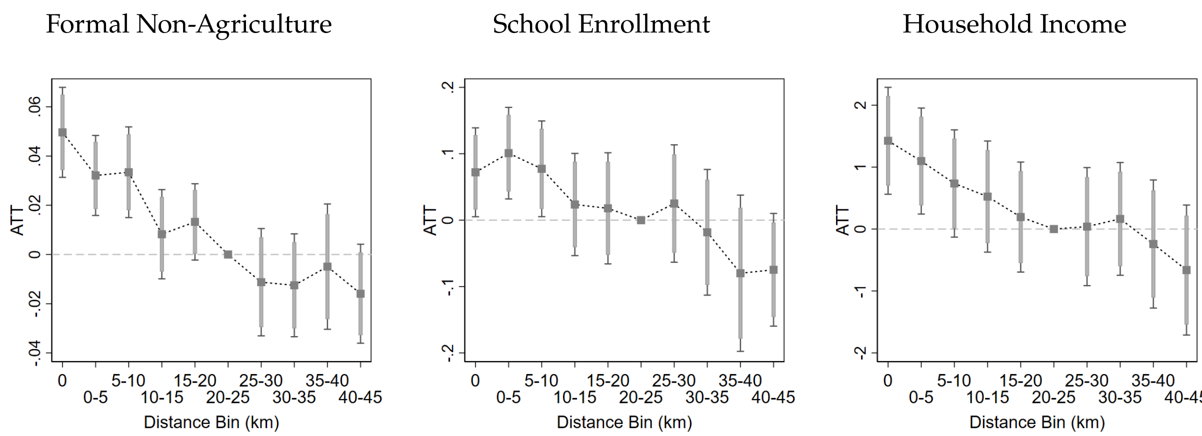
Importantly, this treatment definition relies on the assumption that there are no spatial spillovers from treated districts (within the 10-km buffer) to neighboring, untreated districts. If spatial spillovers exist, meaning that the outcomes in untreated districts are affected by their proximity to treated 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 empirically assess the validity of this treatment definition and to further test the SUTVA assumption, I follow the existing literature (e.g., Gallé et al., 2024) and conduct a distance-based analysis using the following specification:

$$y_{dt} = \gamma_d + \gamma_t + \sum_{b=0, b\neq 5}^{9} \delta_b \text{POST}_{dt} \times \text{Distance}_{d=b} + \theta \text{POST}_{dt} + \varepsilon_{dt} \tag{C1}$$

In this equation, y_{idt} represents the outcome of interest for individual or household i in district d in year t. The terms γ_d and γ_t represent district FE and year FE, respectively. The variable POST_{dt} is a binary indicator for the post-treatment period, taking a value of 1 for post-treatment years and 0 for pre-treatment years for district d. The variable $\text{Distance}_{d=b}$ is an indicator variable that indicates whether district d falls within distance bin b from the industrial zone. The distance bins are defined as follows: $b = 0$ represents the zone-hosting district, $b = 1$ corresponds to districts within 0-5 km from the zone, $b = 2$ represents districts within 5-10 km, $b = 3$ for districts within 10-15 km, and so on, in 5 km increments up to 45 km for $b = 9$. In this model, the coefficient δ_b for $b \in \{0, 1, 2, ..., 9\}$ captures the treatment effects for districts in each of the distance bins, relative to the distance bin $b = 5$ (20-25 km).[7] If we were willing to assume that the reference districts in 20–25 km distance are unaffected by the zones, then δ_b can be interpreted as the effect of this policy on the treated districts in bin b.

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



Notes: Dark shaded areas represent 95% confidence intervals, light shaded areas represnts 90% confidence intervals, diamond symbols represent estimated coefficients on POST_b for bin b from equation (C1). The outcome is employment share in formal non-agriculture for workers 25–64 without a high school diploma (left), schooling enrollment for individuals 15–24 without a high school diploma (middle), and total household income across all sectors (right). Sampling weights are applied throughout.

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, youth school enrollment, and total household income, the coefficients are positive and generally statistically significant within the 0–10 km range. However, these effects quickly diminish with each successive distance bin and generally become statistically indistinguishable from zero beyond 10 km. These findings highlight the localized nature of the benefits associated with proximity to industrial zones,

[7] Because a district may fall within different distance bins for different industrial zones, I assign each district to the distance bin corresponding to the minimum distance to the nearest industrial zone. Additionally, I use the year of establishment of this closest zone as the cutoff to determine the post-treatment period for that district.

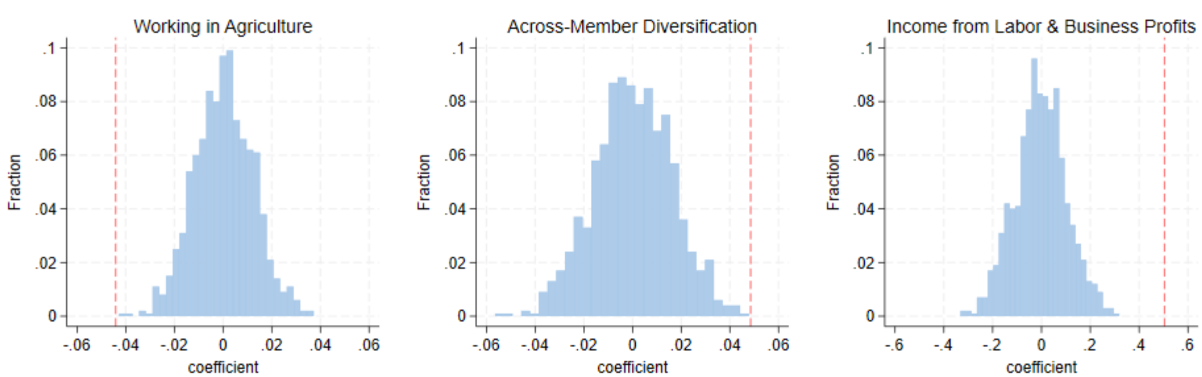
reinforcing the existing literature that documents how the economic and social spillovers of place-based policies tend to be highly concentrated around the immediate vicinity of the zones (Abagna et al., 2025; Ehrlich & Seidel, 2018; Gallé et al., 2024; Tafese et al., 2025). Notably, for the employment outcome, point estimates in the 10–20 km range remain positive, though statistically insignificant. If anything, this suggests that using the 10-km radius as the treatment cutoff likely yields conservative estimates, as districts experiencing modest positive spillovers are included in the control group.

C.8 Monte-Carlo Permutation Test

As an additional check on the econometric specifications, I conduct a placebo test with Monte Carlo analyses of equation (7) to ensure that the current approach provide correct inference and unbiased estimates. Specifically, in each Monte Carlo iteration, I randomly reassign the treatment status series from one district to another district's individual-level labor market data and household-level data, and then test for the impact of industrial zones. The idea is that incorrect assignment of treatment status should yield results of smaller magnitude with zero mean or different sign and statistical insignificance.

Figure C2: Placebo Test: The Impact of Industrial Zones

Results from Monte Carlo Permutation Tests



Notes: Each panel shows the distribution of estimated coefficients (blue bars) for the average cumulative (total) effects of industrial zones on three outcomes: employment in agriculture among 25–64 individuals without a high school diploma (left), across-member labor diversification among households without a high school graduate (middle), and household income from labor and business profits for households without a high school graduate (right). These estimates are generated from 1,000 Monte Carlo simulations in which the treatment status time series of one district is randomly reassigned to another district's labor and household data. The red vertical line represent the baseline estimates. Sampling weights are applied throughout.

Figure C2 presents the distribution of the estimated coefficients with random assignment for three main outcomes: employment in agriculture among individuals aged 25–64 (left), across-member labor diversification measure among households without a high school graduate (middle), and household income from labor and business profits across all households (right). The set of baseline estimates fall outside the resulting distribution of spurious random reassignment estimates—which are centered around zero, suggesting that the industrial zones impact is unlikely to arise by chance. Furthermore, the observed Type I error rates across all outcomes are approximately 5-6% when evaluating at the 5% significance level. These findings suggest that the inference is fairly accurate against the null hypothesis of no industrial zone effect.

C.9 Leave-One-Out Test

To test whether the results are driven by any single district, I perform a leave-one-district-out exercise. By sequentially excluding all observations from one district at a time and re-estimating the model, this approach captures the extent to which local idiosyncrasies drive the overall findings. The results in Figure C3 show that the estimated coefficients remain relatively stable across iterations, with most estimates being close to the baseline values. This consistency suggests that the observed effects of industrial zones on agricultural employment, household labor diversification, and income from labor and business profits are not driven by outlier districts.

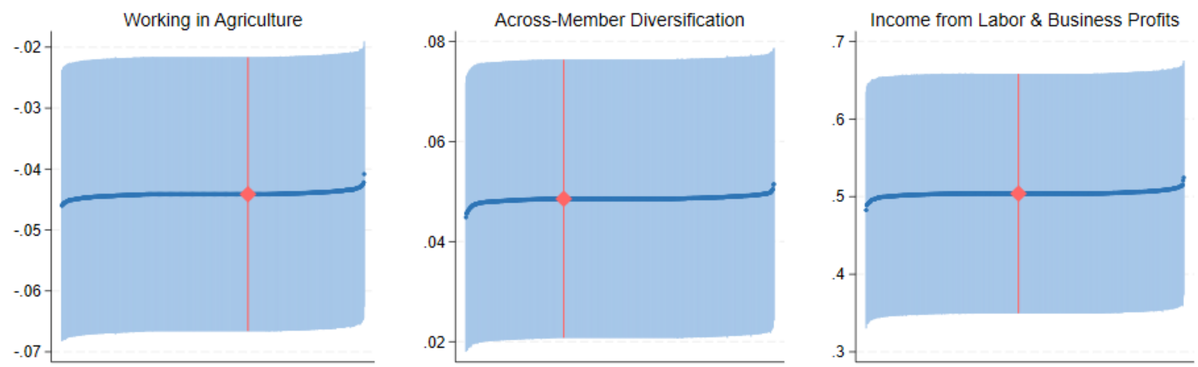
C.10 Province-level and Household Panel Analysis

A potential concern with the baseline analysis is that the VHLSS is designed to be representative at the province level, not the district level, so district-level estimates may be affected by compositional changes or sampling variation. While controlling for demographic characteristics and restricting the sample to long-term residents does not materially change the results (Appendix Table C1), I provide additional evidence using province-level data.

Analyzing province-level effects presents two challenges. First, most provinces have multiple industrial zones established over the study period—some with many zones predating the study—so the approach used in the district-level analysis is not appropriate. Second, the smaller number of observations at the province level (63 provinces) raises concerns about statistical power. With these trade-offs in mind, I estimate event-study effects following the heterogeneity-robust

Figure C3: Sensitivity Test: The Impact of Industrial Zones

Results from Leave-One-District-Out Exercises



Notes: Each panel displays the distribution of estimated coefficients (blue circles) and their 95% confidence intervals for the average cumulative (total) effects of industrial zones on three outcomes: employment in agriculture among 25–64 individuals (left), across-member labor diversification among households without a high school graduate (middle), and household income from labor and business profits (right). The estimates are derived from a leave-one-district-out exercise, in which all observations from a given district are randomly excluded from the sample. Red diamonds indicate the baseline (observed) estimates (Table 3 Column 1, Table 4 Column 1, Table 5 Columns 4/8), and red bars show the corresponding 95% confidence intervals. Sampling weights are applied throughout.

difference-in-differences framework of de Chaisemartin and d'Haultfoeuille (2024) and de Chaisemartin et al. (2024). The treatment variable is the cumulative number of industrial zones in each province. Let F_p denote the first period when province p's number of zones changes from its baseline value. The estimator DID_ℓ compares the outcome evolution of provinces experiencing treatment changes ("switchers") to provinces with the same baseline number of zones whose treatment has not yet changed, capturing the average effect of having been exposed to a higher number of zones for ℓ periods. Standard errors are clustered at the province level, and sampling weights are applied throughout. Estimation is implemented using the Stata package `did_multiplegt_dyn`.

The resulting estimates capture the average effect of one additional zone across all households in the province, including those far from industrial zones, and are thus expected to be smaller in magnitude than district-level estimates. As shown in Table C2, province-level estimates are qualitatively consistent with the district-level findings. The effect on agricultural employment is attenuated from -0.057 to -0.024, while the effect on income from labor and business profits falls from 0.540 to 0.335. Both remain statistically significant. The effect on across-member diversification is similar in magnitude (0.048 vs. 0.039) but loses precision at the province level, likely due to reduced statistical power.

Table C2: Industrial Zones: District-level and Province-level Analyses

	Working in Agriculture*	Across-Member Diversification	Income from Labor and Business Profits
	(1)	(2)	(3)
Baseline result	-0.056	0.049	0.504
	(0.011)	(0.014)	(0.079)
Province analysis	-0.034	0.039	0.400
	(0.016)	(0.052)	(0.085)
Household panel analysis	-0.039	0.071	0.444
	(0.026)	(0.026)	(0.168)

Notes: This table compares district-level and province-level estimates of industrial zone effects. The baseline results (Table 3 Column 1, Table 4 Column 1, Table 5 Columns 4) use a binary treatment indicator at the district level. Province-level estimates use the number of zones established as a continuous treatment variable, following the event-study specification for continuous treatments proposed by de Chaisemartin and d'Haultfoeuille (2024) and de Chaisemartin et al. (2024). For household-panel analysis, "working in agriculture" is measured as the share of hours working in this sector across all working members of the households. Standard errors, clustered at the district level for baseline results and at the province level for province-level analysis, are shown in parentheses. Sampling weights are applied throughout.

Another advantage of the VHLSS is its rotating panel structure, which allows households to be followed for up to three consecutive survey waves. Exploiting this feature, I construct a household panel by pooling four three-wave panels: 2004–2006–2008, 2010–2012–2014, 2012–2014–2016, and 2014–2016–2018. I then re-estimate the baseline specification in equation 7 on this sample to assess the robustness of the baseline results. Note that for this analysis, "working in agriculture" is measured as the share of hours working in this sector across all working members of the households. The findings are qualitatively similar, albeit less precisely estimated, indicating that changes in household demographic composition are unlikely to drive the main results.

Appendix References

Abagna, M. A., Hornok, C., & Mulyukova, A. (2025). Place-based policies and household wealth in africa. *Journal of Development Economics*, *176*, 103482.

Baker, A., Callaway, B., Cunningham, S., Goodman-Bacon, A., & Sant'Anna, P. H. (2025). Difference-in-differences designs: A practitioner's guide. *Journal of Economic Literature*.

Balassa, B. (1964). The purchasing-power parity doctrine: A reappraisal. *Journal of Political Economy*, *72*(6), 584–596.

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., Ciccia, D., D'Haultfœuille, 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. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

Demombynes, G., & Vu, L. H. (2016). *Vietnam's household registration system (english*. World Bank Publications. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/158711468188364218

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

Moretti, E. (2010). Local multipliers. *American Economic Review*, *100*(2), 373–377.

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.