G²LM|LIC
Gender, Growth, and Labour Markets in Low Income Countries Programme

G²LM|LIC Working Paper No. 114 | August 2026

Blossoming Towns

Menna Bishop (University of Warwick)
Céline Zipfel (Stockholm School of Economics)
Niclas Moneke (University of Oxford)

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid from the British people

G²LM|LIC Working Paper No. 114 | August 2026

Blossoming Towns

Menna Bishop (University of Warwick)
Niclas Moneke (University of Oxford)
Céline Zipfel (Stockholm School of Economics)

G²LM|LIC
c/o IZA – Institute of Labor Economics
Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-510

Email: glm-contact@iza.org

IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

UKaid from the British people

G²LM|LIC Working Paper No. 114 | August 2026

ABSTRACT

Blossoming Towns*

Towns and cities create jobs and income. Can the arrival of new jobs also create a town? We study large agro-industrial labour demand shocks in rural sub-Saharan Africa: flower farms in Kenya. Rising production costs in Europe generate a boom in Kenyan cut-flower export production, providing stable wage jobs for mostly low-skilled women. We exploit the specific requirements of cut-flower production to track the arrival and growth of greenhouses over the last two decades. We find large increases in wage employment and home-to-market transformation within agriculture, especially for women. Urbanisation unfolds: population growth and in-migration are flanked by an emerging private rental market for housing. Housing quality and infrastructure access improve, wealth accumulates. Educational investments increase while fertility decreases. Suggestive evidence highlights that the resulting towns initiate occupational change, motivate forward-looking investment and even survive flower farm closures – indicating towns that can sustain themselves as centres of economic activity.

JEL Classification:

J21, O14, O18, R11, Q13

Keywords:

Town formation; wage employment; agro-industry; Kenya

Corresponding author:

Menna Bishop (University of Warwick)

Email: menna.bishop@warwick.ac.uk

* We thank Martina Bj¨orkman Nyqvist, Oriana Bandiera, Gharad Bryan, Dennis Egger, Torsten Figueiredo Walter, Doug Gollin, Morgan Hardy, Jonas Hjort, Peter Hull, Supreet Kaur, Rocco Macchi-avello, Ameet Morjaria, David Nagy, Tanner Regan, Abhijeet Singh, Chris Woodruff, and audiences at AFES Abidjan, CSAE, EBRD, FILMAD Siracusa, Frisch Centre, Gothenburg, Linnaeus, NYUAD, G2LM|LIC, Stockholm School of Economics, KIEL-CEPR African Economic Development Conference, STEG Agriculture, Economic Growth, and Structural Transformation workshop, Nordic Conference in Development Economics, International Conference in Development Economics, Warwick, UEA George-town, and UEA Barcelona for helpful suggestions and comments. Molly Bartlett, Gariyasi Garg, Marcus Harvey Isaksson, and Siyu Li provided outstanding research assistance. Financial support from STEG, G2LM|LIC and the Handelsbanken Foundation is gratefully acknowledged.

1 Introduction

Towns and cities are engines of economic growth and development. Their high population density creates jobs and raises incomes. Can the arrival of well-paid jobs in rural, low-density locations also generate a town?

Rural sub-Saharan Africa is rich in population, yet this population and the economic activity it generates are widely dispersed across space. Low density follows naturally from the dominance of smallholder agriculture (Gollin and Rogerson, 2014), but deprives villages of the agglomeration externalities well-documented in towns and cities (Ciccone and Hall, 1996; Glaeser and Gottlieb, 2009; Ahlfeldt and Pietrostefani, 2019), including in low-income settings (Chauvin et al., 2017; Harari, 2020; Henderson et al., 2021). Path dependence can allow an agglomeration to self-sustain once initiated (Bleakley and Lin, 2012), even in the face of adverse shocks (Davis and Weinstein, 2002).

Urban employment is skewed towards non-agricultural jobs, which are more productive than agricultural ones (Gollin et al., 2014). Therefore, moving rural workers to urban locations has received considerable attention (Bryan and Morten, 2019; Lagakos, 2020). In contrast, our knowledge of how rural locations can transform into towns remains limited.[1] This is despite even small towns seemingly providing many of the benefits of urbanisation, including higher income and consumption (OECD, 2022).[2]

In this paper, we ask whether the arrival of jobs in a rural location can aggregate economic activity and create a town. We study large agro-industrial labour demand shocks that create and concentrate thousands of 'good' jobs – in the sense of providing stable wage employment – in previously rural, agrarian locations in sub-Saharan Africa.

Our context is the arrival and expansion of flower farms in the Kenyan highlands over the last two decades. Driven by sustained demand for cut flowers and rising labour, energy and fertilizer costs in Europe, Kenya experienced a prolonged boom in the production of cut flowers for export to European markets. We construct a geo-identified dataset tracking the entry and expansion of 238 flower farms since 2000. We exploit the peculiar agroclimatic, geological and economic requirements of flower production to identify locations that were similarly suitable for floriculture before the boom. The high resolution and universal coverage of three geo-identified censuses (1999, 2009 and 2019) allow us to analyse the effects of flower farm entry using rich microdata at the level of stable spatial units as small as villages.

We provide six findings: first, flower farm entry increases wage employment, especially for women. Second, population increases sharply, driven by migrants. Third, urbanisation unfolds, with private rental markets, improved quality of housing and

[1] A notable exception is Nagy (2023) who studies endogenous formation of cities from market centres.

[2] Using micro data from sub-Saharan Africa, OECD (2022) show that differences in many household-level outcomes between rural areas and small towns (<50,000 people) are similar to those between small towns and large cities (>1m people), as the former already provide valuable features, e.g., infrastructure.

infrastructure access, all starting to resemble Kenyan cities. Fourth, household welfare increases. Fifth, female demographic choices align with those of their urban counterparts, with younger women giving birth to fewer children, and older women becoming more likely to divorce. Finally, we provide suggestive evidence on occupational diversification, forward-looking capital investments, and the experience of towns after flower farm closure – providing a coherent picture that these towns appear likely to persist. Our findings point to self-sustaining centres of economic activity – blossoming towns – that cause large, long-run transformations of the local economy.

Figure 1: Satellite imagery of a Kenyan settlement before and after the opening of a flower farm.

(a) 2003 **(b)** 2013 **(c)** 2024

Notes: Google Earth Pro satellite imagery of a settlement in rural Kenya, shown before and after the arrival of a flower farm, indicated by the highly reflective silver surface of its greenhouses which first appear in late 2010. Panel (a) depicts the area in 2003, prior to entry; panels (b) and (c) show the same location in 2013 and 2024, respectively, after entry. Original settlement centroid located approx. 3km to the southeast of the flower farm. Over time, new buildings appear and the settlement's spatial footprint expands, transforming into a small town. Original colours minimally edited to align across years and aid visual comparison.

Figure 1 provides a visual summary of the natural experiment at the heart of this paper. Before flower farm entry in 2010, panel (a) shows a small village surrounded by smallholder land. Three years after entry, panel (b) shows more than 25 greenhouses around 3km north of the village, alongside substantial village growth, new streets, and a widened district road. A decade later, the farm has expanded to more than 40 greenhouses and the settlement has grown further and is markedly denser, particularly along the main road, with visible features characteristic of urban form (Henderson et al., 2021; Henderson and Liu, 2023).[3] We interpret this example as *prima facie* evidence that the arrival and concentration of wage jobs can have sizeable long-run effects on urbanisation and densification in a previously rural low-income setting.

That towns could emerge when factories bring jobs to rural areas resonates with historical evidence from industrialising Europe (Marshall, 1890). During the Industrial Revolution, textile factories were established in rural locations in Lancashire and

[3]Zooming in further reveals businesses, an expanded school, a new health facility, and a minibus station. Zooming out, further expansion of the town towards the east and south becomes visible.

Yorkshire, employing large numbers of women and transforming their surroundings into 'mill towns' (Wadsworth and Mann, 1931; Ashmore, 1982). Many remain towns or cities today despite the collapse of the British textile industry, although Heblich et al. (2026) show that the more diversified towns enjoyed greater long-run productivity. Our context provides a modern analogue, allowing us to ask whether large-scale labour demand shocks can create self-sustaining, diversified centres of economic activity.

New flower farms in rural Kenya concentrate a relatively large labour force in one location, providing a local labour demand shock: the median (mean) flower farm in 2010 covers 18 (31) hectares, and employs approximately 317 (539) workers, preferably low-skilled women (Kuiper, 2019).[4] Given export market labour regulation, workers are hired on permanent wage contracts, which are rare in the rural economy.[5]

To produce cut flowers such as roses, flower farms combine low-skilled labour with land for greenhouses, sunlight for photosynthesis, water for irrigation and soil for nutrients to grow flowers. Production involves labour-intensive manual activities such as planting, fertilising, cutting, packaging and transporting cut flowers, supported by light machinery and initial capital outlay for land and greenhouse structures.[6] To reap scale economies in production, flower farms organise in dense compounds resembling factories (Hjort, 2014; Macchiavello and Morjaria, 2015; Ksoll et al., 2023). Ethnographic field work characterises flower farm jobs as agro-industrial in nature (Kuiper, 2019).

Importantly, flower farms provide more than just a large labour demand shock to the local economy: farms also actively attract migrant workers once the local labour pool is exhausted, and commonly contribute wider investments in public services and infrastructure. This may include funding for health facilities, schools, nurseries, and worker accommodation, as well as for bus stops, piped water, electricity and small road upgrades. Our analysis captures the combination of all of these dimensions of flower farm entry – a 'big push' in the sense of Rosenstein-Rodan (1943).

Therefore, our setting lies at the intersection of theories of new economic geography and industrialisation. Both emphasise increasing returns that can move an economy between equilibria (Murphy et al., 1989; Krugman, 1991, 1995). In our setting, flower farms require a large upfront investment to open, while export to high-income countries generates the returns needed to pay workers a wage premium. Hence, flower farms appear to create the conditions for a local high density, high (proto-)industrialisation equilibrium after entry, which we test empirically.

To this end, we construct a spatial dataset tracking the universe of visible Kenyan flower farms from 1999 to 2023. Historical satellite imagery allows us to date farm

[4] Based on benchmark production figures of 17.6 workers/hectare shared by Kenya Flower Council.

[5] In the 1999 census, 20.7% of working-age adults in Kenya report working for pay for someone else, almost exclusively confined to the 22.3% of households living in formally 'urban' settlements in 1999.

[6] Given its perishable nature, the final product requires cooling and rapid onward transport by refrigerated truck to an airport, usually Jomo Kenyatta International Airport in Nairobi, for export by air.

entry and expansion. We link these data to three census rounds, in 1999, 2009, and 2019. The censuses delineate Kenya into roughly 6,000 sublocations, giving us unusually fine spatial detail over more than two decades. They allow us to track employment, urban status, housing and rental markets, infrastructure and household wealth, education, and family formation. We complement these with annual nighttime lights and a geo-identified endline household survey. Together, these data let us follow both the arrival of flower farms and the long-run transformation of surrounding settlements.

Flower farm location choices are not exogenous: cut flowers require specific agroclimatic conditions that make farms concentrate in the moderate climate and water-rich environment of the Kenyan highlands.[7] While benefiting from altitude and abundant water for irrigation, farms also require flat contiguous plots to reap scale economies from greenhouse cultivation. Access to infrastructure for rapid export via cold-chain air freight is also critical. We exploit this complex combination of location factors to construct several dimensions of locational suitability to grow cut flowers. Together, these siting factors provide us with a suitability score to match flower farm locations to plausible control locations that were equally likely to receive a flower farm at baseline. We report placebo specifications to test for pre-trends for all outcomes throughout the paper, and show robustness of all main results to individual dimensions of suitability.

We exploit the staggered arrival of flower farms: combining farm entry with three census rounds gives us two treatment cohorts, 84 sublocations first treated between 2000 and 2009 and 55 first treated between 2010 and 2019. We estimate group-time average treatment effects following Callaway and Sant'Anna (2021), using the doubly robust estimator of Sant'Anna and Zhao (2020) to condition on floriculture suitability.[8]

We find that flower farm entry changes both how people work and where they live. Women's wage employment rises by 12% relative to the never treated at baseline. Work on family farms and businesses falls by an equivalent amount. Population also grows substantially around the farms: the adult population rises by around 970 people in the first census after entry, or 38% of the never-treated baseline, and by 3,510 one census later, more than doubling the baseline. In-migration accounts for a large part of this growth. Recent in-migrants increase by around 530 adults in the first census, 75% and 63% relative to their respective baselines, and by around 2,010 adults one census later.

As population grows around flower farms, the settlements themselves become more urban. In the first census after entry, the share of households living in neighbourhoods classified as urban rises by around 5 percentage points and the share renting from a private landlord by 3 points. Housing quality and household wealth also improve, by 0.13 and 0.16 standard deviations respectively. Annual nighttime luminosity rises after

[7]These growing conditions are not unique to Kenya: highland locations in Ethiopia, Zambia and Colombia have also attracted cut-flower production for export to high-income countries.

[8]We exclude sublocations already exposed to a flower farm in 1999 since we lack pre-treatment data.

farm entry and continues to increase over the following decade, reflecting increases in economic activity and/or electrification progress. Secondary educational attainment rises, especially for adolescent girls. Fertility decreases and family formation norms appear to converge to those seen in Kenyan cities, especially around divorce.

We then ask whether this densification generates activity beyond the farms themselves. First, cross-sectional comparisons in the 2019 census show lower agricultural employment shares near flower farms, offset by gains in trade, services, transport and construction. Second, the 2024 geo-identified household survey shows greater investment in land, housing and business equipment, and less in livestock. Such investments tie people and capital to a location and may help sustain local activity. Finally, we examine persistence more directly through the closure of two large flower farms. Nighttime luminosity in the flower-farm town initially fell but recovered quickly, while the town itself showed no visible contraction. These patterns suggest that the initial concentration of workers can generate persistent towns.

This paper contributes to three strands of literature. First, we contribute to the literature on cities as persistent engines of economic growth due to agglomeration externalities (Ciccone and Hall, 1996; Glaeser and Gottlieb, 2009; Ahlfeldt and Pietrostefani, 2019), including in low-income countries (Chauvin et al., 2017; Harari, 2020; Henderson et al., 2021). Whereas Bleakley and Lin (2012) highlight the persistence of existing towns, we study how new towns form.[9] Our setting provides a modern counterpart to the historical emergence of 'mill towns' around factories in today's high-income countries (Wadsworth and Mann, 1931; Hoskins, 1955; Ashmore, 1982). Heblich et al. (2026) study how concentrated labour demand transformed landscapes previously characterised by relatively even population distributions into such towns, tracing their emergence and eventual decline. Flower farms similarly concentrate large numbers of workers in factory-like production sites in an initially rural landscape, despite remaining formally within agriculture (Hjort, 2014; Macchiavello and Morjaria, 2015; Kuiper, 2019; Ksoll et al., 2023). We show that this labour demand shock generates sustained population growth and densification. This also complements evidence on the returns to density within existing cities (Harari, 2020) by studying the opposite end of the settlement-size distribution: the formation of small towns from dispersed rural populations.[10] Flower farm entry may thus push settlements towards the 'intermediate' initial density associated with subsequent population growth (Michaels et al., 2012).

[9]We also confirm the seminal finding by Davis and Weinstein (2002) in our context that location fundamentals, such as benign agroclimatic conditions for export crop cultivation, and increasing returns to scale, such as the entry cost of factory-like flower farm operations, are crucial determinants of the location and degree of concentration of economic activity.

[10]Land market frictions, as evidenced by Acampora et al. (2025) in rural Kenya, may impede such organic town formation in the absence of external shocks.

We contribute to an emerging second strand of literature on structural transformation in sub-Saharan Africa (SSA). Recent work broadens the concept of structural transformation in this context (Gollin and Kaboski, 2023) beyond the classic reallocation of workers from agriculture to manufacturing (Herrendorf et al., 2014; Gollin et al., 2014) to include rural-to-urban (Bryan and Morten, 2019; Lagakos, 2020), home-to-market (Dinkelman and Ngai, 2022), and self-to-wage employment transformations (Bandiera et al., 2022). Our context combines these dimensions: rural locations become denser and more urban, while workers, particularly women, move into market and wage employment. Yet standard sectoral measures miss this transformation because flower farm workers remain classified in agriculture. Since wage employment remains rare in rural SSA (Bandiera et al., 2022), existing work has focused on moving rural workers to urban locations (Bryan and Morten, 2019; Lagakos, 2020). We instead study how rural workers can concentrate in an initially rural location, forming a denser settlement. Therefore, we also add to recent descriptive evidence on densification in SSA (Bryan et al., 2025) which documents widespread proto-urbanisation in rural areas, often driven by settlements *just* breaching urban density thresholds – akin to our flower farm towns.[11] More broadly, our results show that moving production to rural areas can generate urbanisation alongside economic growth, consistent with flower farm towns developing as production rather than consumption cities (Gollin et al., 2016).

Third, our findings speak to the literature on 'big push' investments and equilibrium selection (Rosenstein-Rodan, 1943; Murphy et al., 1989). Flower farm entry also comes with investments beyond the jobs themselves. Farms face substantial fixed costs, attract a large pool of labour that may benefit other firms and downstream industries, and invest in local infrastructure and services, leaving ample scope for complementarities and increasing returns (Kline and Moretti, 2014; Moneke, 2023). Consistent with this, we find that farm entry attracts large numbers of migrant workers and is followed by broader occupational diversification and improvements in household welfare. This distinguishes our mechanism from the local general-equilibrium effects of large cash transfers (Egger et al., 2022): flower farms generate earned income while drawing workers into the local economy, making migration an important part of the response. Finally, we provide suggestive evidence that the resulting towns persist beyond the initial flower farm (Bleakley and Lin, 2012). Overall, our results suggest that concentrated investments can create new, potentially self-sustaining economic centres, with sizeable local development effects.

[11] Despite their rural location, these towns defy their natural role as agrarian markets in the urban hierarchy (Henderson et al., 2025), instead resembling manufacturing towns that export final products.

2 Setting: Kenya's cut-flower boom

Kenya, a lower-middle-income country with a GDP per capita of $5,842 in 2024 (2021 PPP), has experienced rapid economic growth and urbanisation, alongside substantial structural transformation, over the past three decades. The share of the population living in urban areas rose from 17% in 1990 to 28% in 2020. While agriculture accounted for roughly half of total employment in 1990, that share had fallen to 33% by 2022 (World Bank, 2026). Kenya's cut-flower industry, one of the country's most dynamic export sectors, expanded rapidly over the same period. We describe its emergence and the labour demand it created below.

2.1 Exports boom to Europe

Kenya's commercial horticulture sector began to develop slowly in the late 1960s and early 1970s, when foreign-owned enterprises established large-scale production in the highland areas around Nairobi and Lake Naivasha (Jaffee, 1992). These locations offered favorable growing conditions and good access to Nairobi's international airport (Button, 2020). One driver of this expansion was the eroding competitiveness of greenhouse production in northern Europe. Rising energy and labour costs, amplified by the 1973 oil crisis, made temperature-controlled winter cultivation increasingly expensive and pushed production toward warmer climates. Israel was among the first to enter European flower markets through open-field and plastic-tunnel cultivation. A second wave of cut-flower offshoring shifted production to African growing locations, especially Kenya (Arvis et al., 2009). The country's equatorial, high-altitude conditions and relatively low labour costs allowed year-round production, while direct air links provided access to European markets.

The late 1990s marked the beginning of a particularly rapid phase of expansion. Figure A1 documents the remarkable boom in Kenya's cut-flower exports during the period of our study: export revenues increased nearly thirty-fold between 1992 and 2021, from less than USD 30 million to over USD 700 million. Over the same period, imports of cut flowers into the European Union, Kenya's principal export market, expanded substantially in both value and volume. The industry's growth also coincided with rising production costs in Europe, including labour, fertiliser, and energy costs.[12]

Today, cut flowers rank among Kenya's most important export products. In 2021, the sector generated USD 766 million in export revenues, making Kenya the world's fourth-largest exporter of cut flowers and the country's second-largest agricultural export after tea (Observatory of Economic Complexity, 2023). The overwhelming

[12]Graphical evidence of these trends is provided in Section S.I of the Online Supplement.

majority of exports continue to be destined for European markets, particularly flower auction markets in the Netherlands.

2.2 Expansion of the sector in Kenya: an agro-industrial labour shock

The expansion of Kenya's floriculture industry generated an unusually large labour demand shock in rural areas. Flower production takes place in greenhouse complexes that combine cultivation, sorting, packing, refrigeration, and logistics facilities. Farms are concentrated in a limited number of regions that combine favorable agro-climatic conditions with access to transport infrastructure and export markets. Early flower production was concentrated around Lake Naivasha (Kuiper and Greiner, 2021). Since 2000, however, flower farms have branched out across the Kenyan highlands.

Unlike many agricultural activities in sub-Saharan Africa, floriculture relies heavily on permanent wage labour. According to data from the Kenya Flower Council (KFC), the median flower farm employs 17.6 workers per hectare. Combining these figures with observed farm sizes in the subset of certified KFC member farms in 2022, this implies approximately 317 jobs per farm at the median cultivation area and approximately 539 jobs at the mean cultivation area. Relative to the scale of surrounding communities, these are large employment shocks: the mean local labour market in our data comprises around 2,500 adults, with many much smaller villages and hamlets. At the national level, the sector accounted for more than 65% of all new wage jobs created in Kenyan agriculture between 2010 and 2015 (Mitullah et al., 2017).

Two features of the industry make it particularly well-suited to study the broader consequences of rural wage employment and urbanisation. First, flower farms predominantly offer stable wage contracts, reflecting the year-round nature of flower production (Mitullah et al., 2017; Kazimierczuk et al., 2018). They therefore provide an alternative to self-employed, quasi-subsistence farming and household enterprises as sources of income for rural households. Qualitative evidence further suggests that earnings in floriculture exceed those in many other horticultural activities (Kuiper and Greiner, 2021). Second, the sector relies heavily on migrant labour. Interviews with industry stakeholders confirm that flower farms commonly draw workers from outside the local area. Historical accounts similarly describe flower farms as recruiting workers from a broad catchment area and transforming previously rural settlements into rapidly growing market towns (Kuiper, 2019; Kuiper and Greiner, 2021).

A further distinctive feature of floriculture is its female-biased workforce. Women account for more than 60% of the workforce and are disproportionately employed in picking, grading, packing, and value-added processing activities (Mitullah et al., 2017). These jobs involve repetitive manual tasks requiring precision and consistency, whereas men are more likely to occupy positions involving machinery operation, chemical spraying, or physically demanding work. Female workers also tend to be concentrated in lower-responsibility production roles than their male counterparts (Dolan et al., 2002). The industry's expansion therefore generated a particularly large increase in local wage employment opportunities for women, with potential consequences beyond the labour market, such as for education and family formation.

3 Data

We combine our newly constructed dataset of the universe of (visible) flower farms in Kenya with several sources of outcome data. Our main source of outcome data is geo-identified micro-data from the 1999, 2009 and 2019 Kenya Population Censuses, which we harmonise at the sublocation level. In Kenya, sublocations represent the sixth (and lowest) administrative level, and the finest geographical unit at which census micro-data can be consistently geo-identified over time. The comprehensive spatial coverage of outcome data from the census is particularly valuable in our setting, where a limited number of geographically dispersed locations receive a flower farm. Employing census data at the sublocation-year level allows us to track impacts on communities in the vicinity of new cut-flower farms around the time of their entry.

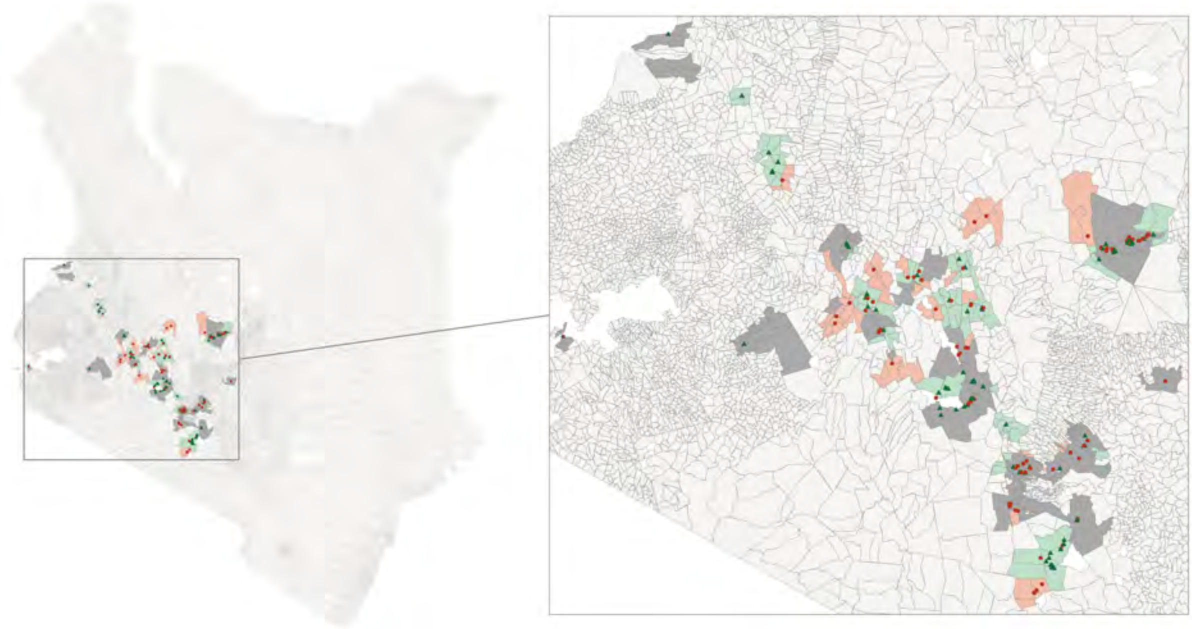
We complement the census panel with two additional sources of outcome data. First, remotely-sensed nighttime luminosity measures provide a proxy for economic development at higher frequency, allowing us to trace the dynamics of local economic activity annually between decadal censuses. Second, a large household survey at endline, the 2024 Kenya FinAccess survey, provides measures of men's and women's welfare that are unobserved in the census – in particular financial inclusion, savings behaviour, and resilience to unexpected financial shocks.

3.1 Universe of (visible) flower farms in Kenya

We build a new database containing the location and timing of entry of flower farms across Kenya for the period 1999-2023. Cut flowers in Kenya are cultivated in greenhouses, which reflect sunlight. This makes greenhouses highly visible from space and easily identifiable by the human eye in remotely sensed imagery – as showcased in Figure 1, panels (b) and (c). Using a combination of Google Earth imagery, LANDSAT imagery, and OpenStreetMap we manually detect, geo-identify, and measure 238 distinct flower farms, concentrated in Western and Central Kenya. In order to obtain the time of entry of each flower farm, we search backwards through historic satellite imagery of the same location, which allows for an approximate determination of when a given greenhouse first appeared. To obtain intensive-margin expansion of flower farms, we also measure the changes in the total area covered by greenhouses in each location in each year that the flower farm (i.e., its greenhouses) was in operation.[13]

In the absence of panel data on flower farm output for the universe of flower farms, we validate our measure of treatment intensity by correlating total area under greenhouse cultivation per year with annual values of cut-flower exports from Kenya.[14] Cut-flower exports follow our proxy for flower farm output remarkably closely over the two decades of our sample.[15] Finally, we compare the distribution of cultivated area in our data with that of a subset of Kenya Flower Council (KFC) member farms: both distributions largely overlap, with a majority of flower farms in both datasets cultivating 10-75 hectares.[16]

Figure 2: Flower farm entry across the Kenyan highlands (2000-2019)



Notes: This map displays the entry of flower farms across Kenya from 2000-2019. All visible greenhouses detected as per the methodology described in Section 3.1 and Appendix A.II. Green dots represent flower farms that opened in the first decade of our sample from 2000-2009, whereas red triangles represent flower farms that opened between 2010-2019. Coloured polygons indicate the treatment timing at the sublocation-level: light grey polygons indicate sublocations that were never within 5km of a greenhouse ('never-treated' sublocations), dark grey ones indicate sublocations excluded from our analysis for one of two reasons: a greenhouse existed within 5km before our study period (159 sublocations), or a greenhouse opened during our study period but at least one census year is missing (29 sublocations). Green (red) polygons represent sublocations that received at least one greenhouse in the 2000-2009 (2010-2019) sample period.

To account for the spatial difference between flower farm locations and the residences of their workforce, we classify a sublocation as treated once at least one flower

[13] Appendix A.II provides graphical examples and additional details on this process.

[14] The rich firm-level survey by Hjort (2014) covers the daily operations of one flower farm over two years (2007-2008). In contrast, a survey of flower farms by Macchiavello and Morjaria (2015) and Ksoll et al. (2023) covers 75 out of 112 identified Kenyan flower farms in one year (2008). Neither suffices to verify the change in scale over two decades of our 238 flower farms, including 55 that only enter post-2009.

[15] Figure A3 provides the time series of greenhouse area and cut-flower exports from customs data.

[16] Further details on KFC member farms are provided in Section S.II of the Online Supplement.

farm starts operating within a set radius of its boundaries. In our main analysis, we use a radius of five kilometres: 139 sublocations receive their first flower farm within five kilometres between 2000 and 2019, including 84 in the first decade and 55 in the second.[17] We exclude 159 earlier-treated sublocations that lie within five kilometres of a flower farm that already existed by 1999, the year of our first census round. A further 29 sublocations are excluded because they are missing outcome data for at least one of the three census rounds. Figure 2 provides an overview of greenhouse location and entry time, as well as treatment time and status across all harmonised sublocations.

3.2 Census micro-data: a sublocation-level panel, 1999–2009–2019

Our main source of outcome data is the 10% random samples of the 1999, 2009 and 2019 Kenya Population Censuses. Since the census covers every inhabited location in the country, these data maximise the spatial overlap between greenhouses and the local communities affected by their arrival. This feature is especially valuable in our setting, where the number of treatment sites is modest relative to the country's size, and nationally representative survey data will likely fail to cover many areas that fall within the relevant local labour markets. To arrive at stable spatial units over time, we harmonise sublocation borders across the three censuses. A small number of sublocations cannot be consistently geo-identified and matched across all three censuses; our balanced panel therefore comprises 5,380 sublocations observed in 1999, 2009 and 2019, out of the more than 6,000 sublocations enumerated.[18] We then match the stable sublocation polygons from the censuses to greenhouse locations described above. Figure 2 displays the resulting set of early-treated (green), late-treated (red), never-treated (light gray) and already-treated or missing (dark gray) sublocations.

We construct the following main outcomes for analysis from the census micro-data:

Employment. We construct outcomes capturing the nature of employment, i.e., whether surveyed individuals are working as wage/salaried workers or are self-employed/unpaid family workers. The wage employment indicator is a binary variable equal to 1 if a respondent reports having worked as an employee for a wage/salary/in-kind payment during the last 7 days, and cites paid employment as their primary economic activity. We construct an analogous family business/farm indicator equal to 1 for respondents who report working on their own – or their family's – agricultural holding or business as their primary economic activity in the past 7 days.[19] All employment outcomes are constructed separately for women and men aged 15-50.

[17] Throughout the paper, we allow for alternative values of the spatial extent of treatment by reporting results for sublocations within one, two and ten kilometre radius from a greenhouse. For reference, the flower farm in Figure 1 is located circa three kilometres away from the settlement that predates its entry.

[18] The estimation sample varies across outcomes from 5,362 to 5,380 sublocations due to missing data.

[19] The two are not mutually exclusive: further responses include apprenticeships and 'other'.

Population and in-migration. We measure local population using counts of women and men aged 15-50, and construct an in-migration indicator equal to 1 if the respondent was born in a different district than the one in which they currently reside. Defining migration at the district-level is in line with the existing (albeit limited) evidence on the origins of migrant labour in Kenya's cut-flower industry. In particular, in a qualitative survey of flower farm workers in Naivasha, the central hub of the industry, Kuiper (2019) finds that a majority of respondents originate from a different province (the highest level of administration in Kenya).[20]

Urbanisation and housing tenure. The census classifies every enumeration area as urban, peri-urban, or rural. We use this official classification to measure the share of urban households in each sublocation.[21] The housing module, available in all three census waves, also records the tenure status of each household's main dwelling. We harmonise these responses into four mutually exclusive categories: owner-occupied housing (whether purchased, self-built, or inherited) and three rental categories, distinguishing between private individual landlords, government landlords (central government, local authorities, and parastatals), and other institutional landlords (private companies, NGOs and faith-based organisations). This distinction provides a rare window into the supply side of local housing markets, allowing us to distinguish market rentals supplied by private landlords from housing provided by employers or the public sector.

Housing quality and household wealth. The housing module also records the physical characteristics of each dwelling, including the materials used for the roof, walls, and floor, as well as the household's main source of drinking water, sanitation, cooking fuel, and lighting. From these variables, we construct harmonised indicators for access to piped water, sewer or septic connections, electricity, and clean cooking fuel. We also code walls, roofs, and floors as finished following the Demographic and Health Surveys' classification, and define a dwelling as finished if at least two of the three materials are finished. We summarise these measures in two principal-component indices, one for housing quality, one for wealth, estimated separately for each census wave within the analysis sample. The former is based on the dwelling characteristics alone. The latter additionally incorporates the household asset variables collected in the 2009 and 2019 censuses (the 1999 index is therefore based on dwelling characteristics only). Both indices are standardised to have mean zero and standard deviation one among control households at baseline, so treatment effects are expressed in baseline standard deviations. All outcomes are aggregated to sublocation-level means.

Children's education. We measure school enrolment separately for girls and boys, distinguishing between primary (ages 6-13) and secondary (ages 14-18) school-age

[20] The 2019 census also records in-migrants' year of arrival, which allows us to compare outcomes of migrants who arrived before and after greenhouse entry within the same sublocation.

[21] As Bryan et al. (2025) show, this *de jure* definition of 'urban' areas tends to be highly conservative in sub-Saharan Africa when compared to alternative measures of *de facto* urbanisation.

children. Our main focus is on secondary enrolment: whereas primary schooling was free and near-universal during most of our study period, secondary schooling remains costly and far from universal. Thus, there is limited scope for local labour demand shocks to increase primary enrolment, while potential increases in income and expected returns to education from flower farm entry could raise secondary enrolment.

Family formation. The censuses collect the 'stock' fertility outcome of the total number of children ever born alive for women of childbearing age, from which we also construct an extensive-margin indicator equal to one if the woman has ever given birth. In addition, we use the month and year of each woman's last live birth to construct a 'flow' measure of recent fertility: an indicator equal to one if she had a live birth within the three years preceding the census. We complement these fertility measures with indicators of marital status: whether the woman is currently married and whether she was ever divorced or separated. We report results for these outcomes separately for women aged 15-29 and 30-44, to distinguish effects during earlier and later stages of family formation.

3.3 Remotely-sensed nighttime luminosity

The decennial frequency of the census conceals the dynamics of local economic activity between census years. We therefore complement the census panel with annual satellite measures of nighttime lights, a widely-used proxy for local economic activity (Henderson et al., 2012). We use two series of annual composites derived from the Visible Infrared Imaging Radiometer Suite (VIIRS): the original VIIRS data available for 2012-2024 at high resolution, as well as the simulated VIIRS dataset that extends the coverage of the VIIRS further into the past using the predecessor nightlights instrument (DMSP-OLS), available at lower resolution for 1992-2023 (Chen et al., 2024).[22]

A concern specific to our setting is that flower farms themselves emit light at night – greenhouses use supplemental lighting, and farm facilities are electrified – so that an increase in measured nighttime luminosity could mechanically reflect the farm's own operations rather than economic activity in the surrounding community. Therefore, our preferred measure of nighttime luminosity masks all pixels within 500 metres of a greenhouse before computing the maximum luminosity within each sublocation's boundaries. Hence, our measure captures changes in light emissions in the area and settlement(s) around the greenhouse, rather than changes originating from the flower farm itself. The annual frequency of these data allows us to estimate event studies for at least eight years before and after greenhouse entry, giving us a sharper test of pre-trends and showing how local economic activity evolves year by year after entry.

[22] Chiovelli et al. (2026) analyse the performance of a related product over the same time period. They highlight beneficial properties of the harmonised time series in detecting changes in economic growth.

3.4 Endline household-level survey on financial access

Finally, we geo-identify the microdata of the 2024 wave of the Kenya National Bureau of Statistics (KNBS) FinAccess household survey, a nationally representative survey of financial inclusion in Kenya. We can geo-identify household responses at the enumeration area-level, which we match to our greenhouses dataset and the associated treated sublocations. The 2024 wave interviews one randomly selected adult in each of 20,871 households across 1,849 sampling clusters nationwide. Matching cluster GPS coordinates to our sublocation shapefile maps respondents into 1,467 sublocations; 16,838 respondents in 1,218 of these sublocations fall within our estimation sample, including 876 respondents in 58 of the 139 treated sublocations.

The FinAccess survey provides rich household- and individual-level measures of household welfare that are rarely observed in routine low-income country survey instruments, and it provides a much higher level of detail compared to the census: among others, individuals are asked about account ownership (bank accounts, mobile money, mobile banking), participation in savings groups ('chamas', Savings and Credit Cooperative Organisations (SACCOs)), saving behaviour, pension contributions, and households' resilience to financial shocks. The FinAccess survey represents a recent innovation by KNBS, and the available microdata permits geo-identification only for the 2024 cross-section. Therefore, FinAccess outcomes do not permit the difference-in-differences research design used in our main analysis; instead, we use them to provide suggestive evidence on local financial access and inclusion in the vicinity of flower farms compared to control locations at endline. However, a highlight of this data is that we can identify a small subset of women who both live in the vicinity of a flower farm, and also report being employed by a flower farm. This allows us to report financial access measures separately for flower farm-employed women, all other women, and men – shedding light on how far economic gains from flower farms spread locally.

4 Empirical strategy

We combine our greenhouse census with the outcome data described above to ask how the arrival of wage employment opportunities affects employment, urbanisation and the broader development of local economies across rural Kenya. Our empirical strategy exploits the staggered arrival of flower farms across locations and relates changes in employment, population, housing markets, living standards, and family formation to greenhouse entry and expansion.

4.1 Staggered arrival of flower farms

We define a sublocation as treated once a flower farm operates within a 5km radius of its boundaries. This distance reflects reasonable commuting distances for flower farm workers, based on consultations with the Kenya Flower Council. However, we also estimate effects at 1, 2, and 10km distance, both to assess sensitivity to the choice of treatment distance and to understand the persistence of effects across space.

Combining the timing of greenhouse entry with the three census waves gives us two treatment cohorts: an early cohort, first treated between 2000 and 2009, and a late cohort, first treated between 2010 and 2019. We drop always-treated sublocations, i.e., those locations that already had a flower farm within 5km at baseline in 1999, for two reasons. First, they cannot contribute to the estimation of treatment effects on the complier population of farms benefiting from the offshoring of production from Europe to Kenya in the 2000s and 2010s. Second, they do not form valid control group locations either since their outcomes are likely to reflect earlier greenhouse exposure, including any effects that persist in the long-run or dynamically continue to evolve after entry (Goodman-Bacon, 2021).

We exploit this staggered expansion in a dynamic difference-in-differences design, estimating group-time average treatment effects following Callaway and Sant'Anna (2021). Identification comes from comparing changes in treated sublocations with changes in locations that have not (yet) received a flower farm. Since flower farms do not locate randomly across Kenya, however, not all such locations provide equally plausible comparisons. We return to this issue in the next section. We aggregate the group-time effects into two post-treatment estimates: the effect at the first census after entry, pooled across the two treatment cohorts in proportion to their number of treated sublocations, and the effect one census later, identified from the early cohort. Standard errors are clustered at the sublocation level.

4.2 Selection into flower farm locations

The main challenge for identification is that flower farm location is not exogenous to locational characteristics in Kenya. Instead, flower farms naturally prefer locations with high productivity in growing flowers that are close to transport links required for export, in particular an international airport due to high perishability of cut flowers. To understand which factors determine flower farm location choices, we conducted numerous qualitative interviews with Kenyan flower farm managers.[23] In addition, we independently verified this information with expert opinions from senior officers at the

[23]These interviews were mostly conducted during the 2023 International Floriculture Trade Fair. Transcripts of all interviews can be found in the Online Supplement.

Kenya Flower Council, who have advised, certified and audited flower farm operations of more than 100 member farms in Kenya since 1996.

These sources agree to a large extent and point to three dimensions of a given location's suitability for floriculture: climate and water, terrain, and infrastructure. These dimensions arise directly from the production function for cut flowers, their perishable nature and the location of demand for the final product in Europe. In particular, they thrive in the mild, moderate climate of the Kenyan highlands, with abundant water resources for irrigation available, coupled with high photovoltaic output. For roses, the cooler highland conditions favor the long stems and large flower heads that command a premium in export markets (Shin et al., 2023). However, the highlands, despite their beneficial climate for cultivation, are geologically rugged, with some areas marked by steep slopes, which prevents efficient cultivation in contiguous rows of greenhouses. Hence, plots of flat land are particularly suitable. In addition, flower farms would ideally require electricity access to extend the life span of flowers once they are cut through refrigeration (IFC, 2020). Finally, rapid transport to end customers is critical via truck (Button, 2020; Ksoll et al., 2023).

These characteristics may also shape local development in the absence of the flower industry, such that an assumption of unconditional parallel trends appears demanding *ex ante*. Therefore, our empirical strategy centres on comparing treated sublocations with untreated ones that were similarly suited to host a flower farm before the flower industry's boom.

To do so, we combine twelve variables capturing the three dimensions above into an overall suitability score, defined as the fitted probability that a sublocation receives a flower farm after 1999. Figure 3a shows the coefficients underlying the score. Elevation and rainfall are the strongest predictors of flower-farm location, consistent with the industry's concentration in the mild, well-watered Kenyan highlands. Road access and baseline electrification also predict farm location, while terrain plays a smaller role. Taken together, climate and water account for roughly four-fifths of the model's explanatory power, with infrastructure and terrain accounting for the remainder.[24] We leave soil nutrients out of the baseline score because they are measured after most farms entered (Hengl et al., 2017) and could themselves have been affected by floriculture. Including soil nutrients barely changes our results (see Appendix A.IX).

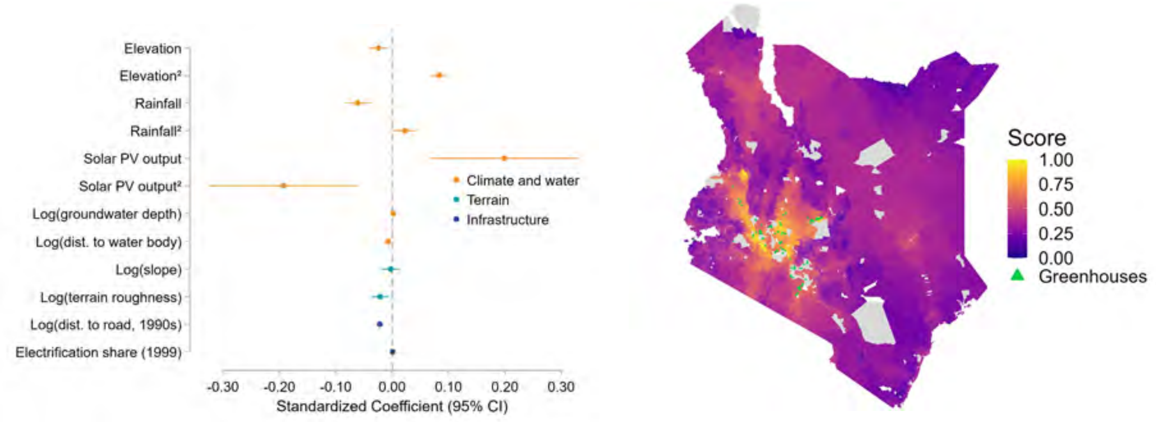
Figure 3b maps the resulting score across Kenya. The pattern closely matches where flower farms actually locate: the most suitable sublocations are concentrated in the highlands, where the industry has developed. Our identifying assumption is therefore

[24] Tables A3 and A4 report the corresponding univariate regressions and the predictive power of the three covariate blocks separately. Elevation is grouped under "climate and water," as it proxies for the mild highland climate particularly suited to floriculture; slope and ruggedness are grouped under "terrain." Our main results use the full suitability score and are therefore unaffected by these groupings. The final column of Table A4 includes all 12 covariates jointly, i.e., the coefficients plotted in Figure 3a.

parallel trends conditional on floriculture suitability. In other words, we compare treated locations with places that were similarly suited to floriculture, and assume that their outcomes would have followed parallel trends had the flower farm not arrived.

Figure 3: Flower-growing suitability score

(a) Components of the suitability score **(b)** Full suitability score



Notes: Panel (a) plots standardized coefficients from the linear probability model underlying the suitability score: an indicator for a greenhouse operating within 5km of the sublocation in 2000–2019 is regressed on the 12 siting components, estimated on the universe of Kenyan sublocations excluding those already within 5km of a greenhouse in 1999. Covariates are standardized to mean zero and unit variance, so coefficient magnitudes are comparable across components; bars are 95% confidence intervals including robust standard errors. Panel (b) maps the fitted score across sublocations, linearly rescaled to the unit interval (the rescaling preserves the ranking and relative spacing of sublocations); greenhouse locations (2000–2019 entrants) are overlaid as triangles. Sublocations already treated by 1999 or with a missing score component are shown in grey.

Our baseline specification uses the group-time framework by Callaway and Sant'Anna (2021) with the doubly robust estimator developed by Sant'Anna and Zhao (2020), and the suitability score entering both the propensity score and the outcome regression. Formally, for each cohort g and each census year $t \geq g$,

$$\mathbb{E}\left[\,Y_t(0) - Y_{g-1}(0) \mid X,\, G_g = 1\,\right] \;=\; \mathbb{E}\left[\,Y_t(0) - Y_{g-1}(0) \mid X,\, D_t = 0\,\right], \tag{1}$$

where X is the suitability score and $D_t = 0$ indicates sublocations not yet treated by year t — a group that always contains the never-treated and, for the early cohort's first post-period, the late cohort as well. Identification also requires overlap: comparisons must exist at every treated level of the score (Heckman et al., 1998; Sant'Anna and Zhao, 2020). Because sublocations far below the treated range tell us little about how treated places would have evolved, we restrict the comparison pool, dropping those whose score falls below the minimum treated score while retaining all treated sublocations (see Appendix A.IV.3). This keeps the effect we estimate defined over all flower-farm locations.[25]

How far can we assess the parallel-trends assumption in the data? Throughout, we test for pre-trends for the late-treated cohort. These sublocations are still untreated between 1999 and 2009, so we use their change over that decade as a placebo and report it alongside the treatment effects as the third row of estimates in all regression tables. However, there are two limits to what we can learn from this. First, for the early cohort, there is no equivalent test. Second, for the late cohort, a flat placebo is reassuring but only covers the decade before entry. Additionally, flower-suitable places could still have followed different trajectories afterwards, for example because of differential investment in the highlands. We therefore prefer to condition on suitability even for outcomes where we fail to detect evidence of pre-trends.

Further details on the construction of the suitability score and our choice of common-support restriction are provided in Appendix A.IV. Appendix A.IX examines the sensitivity of the results to these and other choices, such as the conditioning set and comparison group, our decision to exclude soil nutrients from the score, the common-support threshold, and alternative estimators and measures of greenhouse exposure.

5 Results

To test whether flower farms concentrate economic activity and form towns, we proceed as follows. We first confirm whether flower-farm entry creates wage employment, and test whether these jobs draw in population to its surrounding area. We then ask whether this concentration of people and economic activity can be characterised as urbanisation, and study the implications of this transformation on the settlement's inhabitants.

5.1 Occupational shifts in the labour market

Table 1 examines how the arrival of flower farms changed the occupational structure of nearby communities. Panel A reports results for women and Panel B for men, both aged 15-50, whom we define as the working-age population. The third row of each panel reports pre-trend estimates for sublocations first exposed to flower farms between 2010 and 2019 (the 'late-treated cohort'). Across employment outcomes, placebo changes between 1999 and 2009 for the late-treated are small and insignificant. This absence of

[25] Under Equation (1) and overlap, each group-time effect is identified by the doubly robust estimand $ATT(g,t) = \mathbb{E}\big[(G_g/\mathbb{E}[G_g] - \omega_{g,t}/\mathbb{E}[\omega_{g,t}])(Y_t - Y_{g-1} - m_{g,t}(X))\big]$, where $\omega_{g,t} = p_{g,t}(X)(1-D_t)/(1-p_{g,t}(X))$ is the propensity odds weight on the not-yet-treated and $m_{g,t}(X) = \mathbb{E}[Y_t - Y_{g-1} \mid X, D_t = 0]$ the counterfactual outcome-change regression. Consistency requires only one of the two working models — the siting propensity or the outcome-change regression — to be correctly specified, not both.

differential pre-trends is reassuring and suggests that any occupational changes emerge during the period in which flower farms arrive, and do not precede them.

Panel A shows that the entry of a flower farm produces a shift towards female wage employment in communities within 5km. The share of working-age women in wage employment rises by 1.2 percentage points (column 2). Relative to wage employment in never-treated areas in 1999, this represents an increase of approximately 12%. Conditional on having worked during the previous seven days, the share of women in wage employment rises by 2.4 percentage points, or around 15% relative to the corresponding never-treated mean. The corresponding estimates remain positive in the following census period for sublocations first exposed between 2000 and 2009 (the 'early-treated cohort'), at 1.4 percentage points for female wage employment and 1.7 percentage points conditional on working, although neither is statistically significant.

The movement into wage employment is accompanied, initially, by a movement away from work in household farms and family businesses (columns 4 and 5). In the first treated period, the share of working-age women engaged in family or farm work falls by 1.8 percentage points, although not statistically significant, while the share falls by 2.4 percentage points among women who worked in the previous seven days.[26]

In line with overall high female labour-force participation in rural Kenya, column 1 shows a more muted effect on overall employment. The probability of having worked in the previous 7 days is essentially unchanged in the first treated period, but rises by 3.8 percentage points in the following decade. The initial increase in female wage employment therefore primarily reflects a reallocation of women across occupations rather than an expansion in overall labour force participation.

The results for men, reported in Panel B, follow a different pattern. In the first treated period, there is little evidence of a broad change in men's employment structure. The probability of having worked rises by 0.6 percentage points, and the unconditional effects on wage employment and family or farm work are small and statistically insignificant. Conditional on working, however, the wage-employment share rises by 2.4 percentage points, while family or farm work falls by 2.5 percentage points. By the following decade, however, employment among men in the early-treated areas has expanded more broadly. The share of men who worked in the previous seven days rises by 5.8 percentage points, the wage-employment share rises by 3.9 percentage points, and the share reporting work in a family business or farm rises by 1.9 percentage points, although the latter estimate is not statistically significant. Conditional on working, wage employment rises by 6.1 percentage points, while the share in family or farm work falls by 6 percentage points. Thus, although the earliest employment response to flower

[26]A decade later, neither estimate is statistically different from zero, which coincides with substantial population inflows into treated areas as discussed below. This implies that changes in employment shares increasingly reflect not only occupational reallocation, but also the arrival of new workers.

Table 1: Employment on flower farm entry

	(1) Work last 7d	(2) Wage job	(3) Wage job (if work)	(4) Self/family work	(5) Self/family work (if w.)
	Panel A: Women				
First census after entry	-0.008	0.012**	0.024**	-0.018	-0.024***
	[0.013]	[0.006]	[0.010]	[0.013]	[0.009]
Second census after entry	0.038**	0.014	0.017	0.025	-0.015
	[0.016]	[0.009]	[0.016]	[0.019]	[0.016]
Decade before entry (placebo)	0.005	0.003	0.007	-0.005	-0.014
	[0.024]	[0.009]	[0.013]	[0.023]	[0.014]
Never-treated mean (1999)	0.668	0.099	0.163	0.555	0.815
Early-treated sublocations	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55
Total sublocations	1,612	1,612	1,612	1,612	1,612
Observations	4,824	4,824	4,817	4,824	4,817
	Panel B: Men				
First census after entry	0.006	0.010	0.024**	-0.005	-0.025**
	[0.010]	[0.009]	[0.011]	[0.011]	[0.012]
Second census after entry	0.058***	0.039***	0.061***	0.019	-0.060***
	[0.013]	[0.013]	[0.016]	[0.017]	[0.016]
Decade before entry (placebo)	0.003	0.011	0.022	-0.014	-0.029
	[0.016]	[0.014]	[0.019]	[0.021]	[0.020]
Never-treated mean (1999)	0.810	0.280	0.353	0.511	0.624
Early-treated sublocations	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55
Total sublocations	1,612	1,612	1,612	1,612	1,612
Observations	4,824	4,824	4,823	4,824	4,823

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates on a balanced panel of sublocations observed in the 1999, 2009 and 2019 censuses. A sublocation is treated once a flower farm operates within 5 km of its boundaries; sublocations already treated in 1999 are excluded. The comparison group pools never-treated with not-yet-treated sublocations. Standard errors, in brackets, are clustered at the sublocation level. "First census after entry" is the average effect at the first census following a flower farm's entry (zero to ten years of exposure), pooled across the 2000–2009 and 2010–2019 entry cohorts; "Second census after entry" is the effect one census (a further decade) on, identified from the 2000–2009 cohort; "Decade before entry (placebo)" is the differential 1999–2009 change for sublocations first treated only in 2010–2019. Doubly robust estimates (Sant'Anna and Zhao, 2020) conditioning on the floriculture suitability score. The comparison pool is restricted to sublocations whose score is at least the minimum score among treated sublocations; every treated sublocation is retained. Wage job = worked for pay in the last seven days. Self/family work = work on the household's own agricultural holding or (non-farm) family business. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

farm arrival is concentrated among women, the longer-run estimates suggest a broader transformation of male employment as the local economy expands. The simultaneous increase in men's overall employment and wage employment is consistent with an expansion of local labour demand rather than occupational reallocation alone.

The census does not consistently ask about individual occupations or activities underlying these changes, which prevents us from directly identifying flower-farm employed workers in census data. However, the main census harmonisation allows us to consistently distinguish between wage work and self-employment (including family employment). In addition, for the 2009 and 2019 censuses, respondents additionally report employer type, allowing us to further unpack the increase in female wage employment. As shown in Table A6, the share of women in the late-treated cohort working for pay in the private sector rises by 1.5 percentage points, and by 2.3 percentage points conditional on working. In contrast, we find no evidence of increased paid work for other employers. The corresponding private-sector estimates for men are positive but less precisely estimated. These results suggest that the increase in female wage employment documented in Table 1 is concentrated in private-sector jobs.[27]

Figure 4 examines how the female wage-employment effect varies with distance from the farms and with the conditioning set. It reports first-treated-period estimates using buffers of 1, 2, 5, and 10km, from unconditional estimates to doubly robust estimates conditioning on terrain, infrastructure, climate and water access, and the full suitability score on the common-support sample. At 2km, the estimates range from about 1.5 to 1.8 percentage points across specifications, and at 5km from about 1.2 to 1.5 points. Pre-trends are close to zero at every distance. The narrow spread of the estimates across conditioning sets suggests that the result is largely robust to accounting for various differences in floriculture suitability.

The figure also shows a spatial pattern that is mirrored across most other census outcomes (see Appendix A.VIII): effects are strongest within 2-5km of flower farms, consistent with accounts of Kenya's flower industry in which workers live in settlements close to the farms (Kuiper, 2019). Estimates are smaller and noisier within 1km, partly reflecting the low number of treated sublocations at this distance. When including sublocations up to 10km away estimates are essentially zero across specifications.

This spatial decay also raises the bar for any confounding effects to explain the non-trivial spatial signature.[28] In combination with the absence of pre-trends at every distance, this supports our interpretation that the effects are caused by flower farms rather than by broader characteristics of the places in which they locate.

5.2 Population dynamics

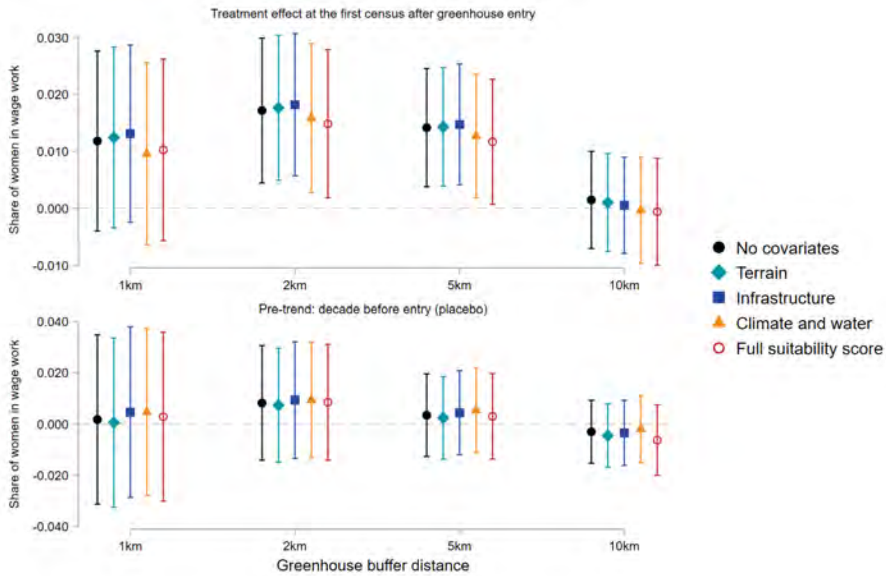
Table 2 shows that the labour market results are accompanied by substantial changes in the size and composition of the local working-age population. Flower farm entry did not merely alter employment among a fixed set of residents. It also attracted people to the surrounding areas, changing the denominator underlying the employment shares.

In the first treated period, the population aged 15-50 in the area around a flower farm increases by 970 persons per sublocation, around 38% of the never-treated 1999 mean

[27] Employer type remains rather imprecise, however, so we cannot distinguish private wage employment on flower farms from similar jobs in the local economy that emerge around the same time.

[28] Reassuringly, we fail to detect evidence of differential improvements in road access after flower-farm entry: Table A5 compares travel time to Nairobi's international airport by flower farm endline status, comparing the 1990s baseline road network with two independent measures of the endline road network. If anything, travel time to the airport improved by less in flower farm locations relative to similarly suitable locations, refuting differential road improvements as an alternative mechanism. Appendix A.V provides details on the construction of these measures.

Figure 4: Female wage labour effects across buffer distances and comparison groups



Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the share of women aged 15-50 in wage employment, by greenhouse buffer distance (1, 2, 5 and 10km); the comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score, a sublocation-level linear-probability prediction of greenhouse presence from all 12 components, whose series drops comparison sublocations with scores below the treated minimum and retains every treated sublocation. All series are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation-level.

of 2,534.[29] The increase is similar for women and men: around 500 (470) additional women (men) of working age. Population density increases by 11% (column 6).[30]

A substantial part of this population growth appears to be driven by in-migration. The number of female residents who had moved into their current county or district during the previous ten years rises by 283 in the first treated period, relative to a never-treated baseline of 379; the corresponding increase for men is 255, relative to a never-treated baseline of 407. Taken together, recent in-migrants account for over half of the total population increase. Since this measure cannot capture within-county or within-district migration and excludes migrants who subsequently moved away again, this represents a lower bound on the contribution of migration to population growth.[31]

The population response grows considerably over time. By the subsequent census period for the early-treated cohort, the working-age population is approximately 3,510 persons higher, around 140% of the never-treated baseline, while population density is

[29] Table 2 counts derived from 10% census samples are scaled by ten to population equivalents.

[30] Given the skewed distribution of population density due to Nairobi, Kenya's largest city, we report effects on log population density. The level effect is statistically significant when excluding Nairobi district from the control group, which also makes all population count results more pronounced.

[31] While we cannot test out-migration directly using census data, industry stakeholders report that farm employment allows working-age residents who would otherwise have left to remain.

around 21% higher. The ten-year in-migrant stock is also 1,056 higher for women and 951 higher for men. Therefore, local population growth represents a sustained process of expansion, instead of a one-off attraction of flower farm workers.

It is important to highlight that the changes in local population are central to the interpretation of Table 1. The employment coefficients are shares of the working-age population or, in the conditional columns, of those currently working. As flower farms attract migrants and expand the local population, these denominators also respond to treatment. A modest change in the wage-employment share can therefore mask a much larger increase in the number of women holding wage jobs. Indeed, the fact that wage employment treatment effects remain positive as the female working-age population expands implies that wage employment increased sufficiently to absorb the incoming population. Hence, the labour market and population estimates should be interpreted jointly: over time, the increase in wage employment shares in Table 1 reflects both the creation of wage opportunities and the changing size and composition of the local population – with in-migration mechanically diluting wage employment shares.

Table 2: Population growth and in-migration on flower farm entry

	(1) Population (total)	(2) Population (women)	(3) Female in-migrants (last 10 years)	(4) Population (men)	(5) Male in-migrants (last 10 years)	(6) Log population density
First census after entry	970.151***	500.044***	282.845***	470.107***	255.111**	0.111***
	[321.429]	[166.794]	[105.947]	[156.116]	[101.694]	[0.033]
Second census after entry	3507.101***	1821.894***	1056.316**	1685.208***	951.106**	0.205***
	[1311.704]	[700.867]	[454.430]	[611.862]	[377.251]	[0.067]
Decade before entry (placebo)	148.584	80.690	-49.169	67.894	-43.208	-0.021
	[355.812]	[182.143]	[56.514]	[175.313]	[56.497]	[0.062]
Never-treated mean (1999)	2533.594	1252.035	379.013	1281.559	406.978	6.637
Early-treated sublocations	84	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55	55
Total sublocations	1,612	1,612	1,612	1,612	1,612	1,612
Observations	4,824	4,824	4,824	4,824	4,824	4,824

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates on a balanced panel of sublocations observed in the 1999, 2009 and 2019 censuses. Treatment definition and specifications as in Table 1. Sublocation-level counts. Col. (1): all persons aged 15–50 (women + men). Cols. (2),(4): women and men aged 15–50. Cols. (3),(5): female and male in-migrants over the last 10 years, i.e. individuals born outside the current county/district whose reported year of arrival falls in the census wave's reference decade (1990–1999 for the 1999 wave; 2000–2009 for 2009; 2010–2019 for 2019). Counts are scaled by a factor of ten to population equivalents: the censuses are 10 percent samples. Col. (6): natural log of the sublocation's census-derived population density (population equivalents per square kilometre). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Although we cannot directly observe whether migrants move to work on flower farms, several pieces of evidence indicate that the inflow of population is closely linked to new employment opportunities created by the sector. Historical accounts describe flower farms as drawing workers from a broad geographic catchment area (Kuiper and Greiner, 2021). Consistent with this interpretation, we exploit information from the 2019 census on both the timing and the stated motive for migration (see Table A7). Restricting attention to in-migrants living in treated sublocations in 2019, we compare individuals who arrived before and after entry of the local flower farm. Migrants arriving after farm entry are significantly more likely to hold wage jobs than earlier arrivals, with particularly large effects among individuals who report work or employment as their reason for moving. While necessarily correlational, these patterns suggest that the expansion of flower farms attracted migrants seeking wage-employment opportunities.

As in Table 1, there is no evidence of differential population trends immediately before treatment. Between 1999 and 2009, areas first treated during 2010–2019 show no statistically significant differential change in population or ten-year in-migration. The point estimates are small relative to the treatment effects: around 150 persons for total population, against an effect at treatment of 970, and around 40-50 in-migrants of each sex, against effects of 255-283. The sharp population growth therefore only arises in the decade of treatment rather than continuing an evident pre-existing trend. The population response is more spatially diffuse than the labour market effects – estimates decline with distance but remain positive and statistically significant even at 10km.[32]

Overall, Tables 1 and 2 document a two-part transformation. Flower farms change the nature of local employment, with women moving from household-based work into private-sector wage jobs. At the same time, the new employment opportunities attract working-age migrants and generate substantial population growth. The resulting communities are not simply villages in which existing residents change occupations; they become larger settlements organized around a new source of formal wage employment.

5.3 Urbanisation

The population inflows documented above should manifest themselves in increased economic activity. While the census only captures information on economic activity every decade, remotely-sensed nighttime luminosity allows us to trace changes at a higher frequency, such as annually (Henderson et al., 2012; Chiovelli et al., 2026).

Figure 5 reports annual event studies using simulation-extended maximum nighttime luminosity (VIIRS) for 1999-2023. No differential pre-trends in luminosity are visible. However, flower farm entry increases luminosity, which rises steadily thereafter, with effects continuing to build over the first decade of exposure. This pattern appears across all four definitions of the spatial extent of the flower farm's catchment area.[33]

Since simulated VIIRS data essentially degrades the resolution of pre-2012 luminosity reflectance, we show robustness to using just the high-resolution, native VIIRS data from 2012 onward. Results are consistent across datasets: no statistically significant pre-trends, and luminosity steadily rising after entry across all catchment areas.[34]

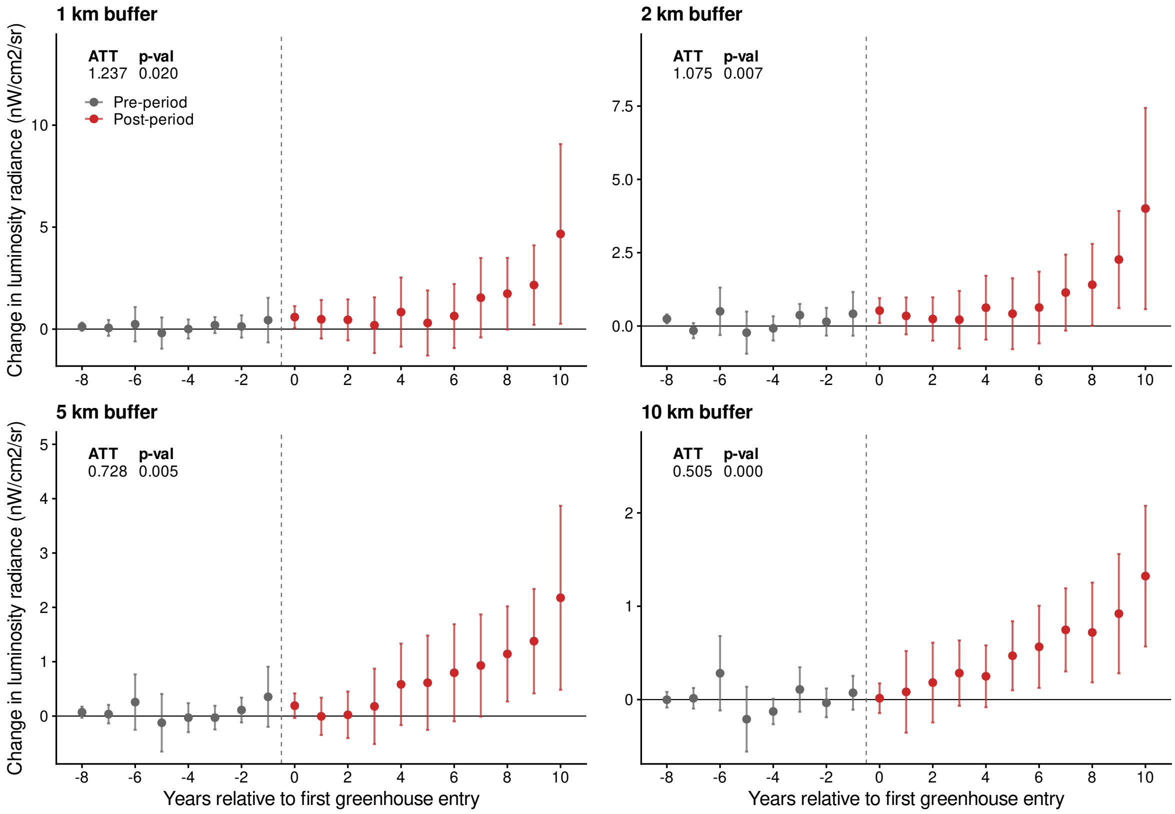
Increases in nighttime lights may partly reflect expanding electrification rather than economic activity alone. The two are hard to separate in this setting, and perhaps

[32]See Figure A7. The infrastructure- and suitability-score specifications generally produce the most conservative estimates, but remain qualitatively similar and preserve the same spatial gradient.

[33]To prevent contamination from greenhouse's own reflectance, we mask all pixels within 500m.

[34]See Appendix Figure A20, where early flower farm entrants are observed only after treatment, resulting in shorter event studies and noisier estimates. Excluding Nairobi leaves results unchanged.

Figure 5: Event studies of maximum nighttime luminosity upon greenhouse arrival (1999-2023)



Notes: Event-study estimates of the effect of flower farm entry on nighttime luminosity (Callaway and Sant'Anna, 2021). The outcome is the maximum nighttime luminosity radiance (nW/cm^2/sr) across pixels within each sublocation's boundaries, computed after masking all pixels within 500 metres of a greenhouse, so that light from the farm's own operations is excluded. Luminosity is taken from the simulated VIIRS time series; sublocations already treated in 1999 are dropped with the estimation window covering 1999-2023. Each panel defines treatment by a different buffer: a sublocation is treated once a flower farm operates within 1, 2, 5 or 10km of its boundaries. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level. Each panel also reports the overall ATT (the average of the post-entry effects over the event window) with its two-sided p-value.

not worth separating: contemporary reports from the Kenya Flower Council describe flower farms as contributors to rural electrification and other local infrastructure. Thus, electrification may represent one plausible channel linking employment growth to wider spatial development (Kenya Flower Council, 2024).[35]

Census data helps unpack this growth in economic activity. Table 3 reports effects on the share of locations officially classified as urban: it rises by 4.8 percentage points initially, and by 11 points later, relative to a never-treated baseline mean of 10.2%.[36]

The tenure results provide insight into how existing settlements absorb the influx of new residents. The share of households owning their dwelling falls by 2.3 percentage points, and the share renting from a private landlord rises by 3.4 points, a 30% increase relative to the never-treated mean of around 12%, which increases further to 7.5 points over time. Renting from government sources declines. Hence, the shift reflects private

[35] Figueiredo Walter and Moneke (2026) show that the effects of rural electrification concentrate in locations with productive users, such as a flower farm, which exhibit externalities on households.

[36] The effect declines almost linearly with distance from the flower farms (see Figure A8a).

Table 3: Urbanisation and housing tenure on flower farm entry.

	(1) Urban	(2) Owner-occupied	(3) Rented: private landlord	(4) Rented: government	(5) Rented: other
First census after entry	0.048**	-0.023**	0.034***	-0.017**	0.005
	[0.022]	[0.010]	[0.010]	[0.007]	[0.005]
Second census after entry	0.110***	-0.057***	0.075***	-0.032***	0.011**
	[0.037]	[0.017]	[0.017]	[0.012]	[0.005]
Decade before entry (placebo)	0.028	0.002	-0.001	-0.005	0.004
	[0.032]	[0.018]	[0.018]	[0.007]	[0.005]
Never-treated mean (1999)	0.102	0.831	0.115	0.047	0.007
Early-treated sublocations	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55
Total sublocations	1,612	1,612	1,612	1,612	1,612
Observations	4,823	4,824	4,824	4,824	4,824

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates on a balanced panel of sublocations observed in the 1999, 2009 and 2019 censuses. Treatment definition and specifications as in Table 1. Household-level shares (one observation per household), averaged to the sublocation. Col. (1): share of urban households. Cols. (2)-(5): tenure of the main dwelling unit, four mutually exclusive groups that sum to one over households with nonmissing tenure. Col. (2) owner-occupied (purchased, constructed, or inherited); Col. (3) rented from a private (individual) landlord; Col. (4) rented from government; Col. (5) rented, other. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

rental markets rather than public housing provision. Intuitively, settlements absorbing large inflows of migrant workers should see rental arrangements expand faster than owner-occupancy.[37] Taken together, the population, urbanisation and satellite evidence suggests that flower farms generate thriving centres of economic activity around them.

If settlements around flower farm locations increasingly resemble towns, one may expect households living in them to increasingly live like town dwellers. To test this, Table 4 reports effects on dwelling and living-standard outcomes. We report both outcomes that measure households' own investment decisions (e.g., finished building materials, clean cooking fuels) as well as networked infrastructure access (e.g., piped water, sewerage and electric lighting). Both types of investments increase significantly.

Households in treated sublocations experience broad improvements in housing quality and utilities. In the first treated period, access to piped water and sewer or septic connections rises by 3.3 and 1.2 percentage points, clean cooking fuel by 3.9 points, and the share of dwellings with at least two finished materials by 4.4 points. The housing and wealth indices rise by 0.13 and 0.16, almost doubling a decade later.

Electric lighting also increases substantially, by 6.9 percentage points in the first treated period and 20.0 points one period later, against a baseline of 7.7%. This outcome already exhibits a positive pre-trend, however, so the effect magnitude may be inflated. Nonetheless, the increase is consistent with luminosity results and anecdotal evidence of

[37] In line with this, Table A8 decomposes the increase by household migrant status. At treatment, essentially all of the 3.4 percentage point rise in the first census comes from households containing a recent in-migrant (column 2), while renting among other households is unchanged (column 3).

flower farms as contributors to rural electrification (Kenya Flower Council, 2024). With the exception of electric lighting, most outcomes in Table 4 show either no differential pre-trend between the late-treated and never-treated sublocations or a negative one.[38]

Table 4: Housing quality and household wealth on flower farm entry.

	(1) Piped water	(2) Sewer/septic connection	(3) Electric lighting	(4) Clean cooking fuel	(5) ≥2 finished materials	(6) Housing index (PCA)	(7) Wealth index
First census after entry	0.033*	0.012**	0.069***	0.039***	0.044***	0.125***	0.163***
	[0.018]	[0.006]	[0.012]	[0.010]	[0.010]	[0.034]	[0.037]
Second census after entry	0.015	0.028**	0.200***	0.114***	0.098***	0.231***	0.252***
	[0.027]	[0.013]	[0.026]	[0.021]	[0.016]	[0.061]	[0.065]
Decade before entry (placebo)	-0.091***	-0.002	0.041***	-0.001	-0.011	-0.018	0.006
	[0.035]	[0.006]	[0.014]	[0.005]	[0.015]	[0.053]	[0.055]
Never-treated mean (1999)	0.197	0.057	0.077	0.029	0.297	0.388	0.388
Early-treated sublocations	84	84	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55	55	55
Total sublocations	1,612	1,612	1,612	1,612	1,612	1,612	1,612
Observations	4,824	4,824	4,824	4,824	4,821	4,824	4,824

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates on a balanced panel of sublocations observed in the 1999, 2009 and 2019 censuses. Treatment definition and specifications as in Table 1. Household-level, one observation per household, then averaged to the sublocation. Col. (1) piped water; Col. (2) sewer or septic connection; Col. (3) electric (grid/mains) lighting; Col. (4) clean cooking fuel; Col. (5) at least two of the dwelling's roof, walls and floor made of finished materials. Cols. (6)-(7) housing and wealth indices: first principal components of dwelling characteristics (the wealth index additionally includes asset ownership), re-estimated within each census wave and standardized to never-treated sublocations in 1999, so estimates read in never-treated standard deviation units. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.4 Education

We now turn to whether flower farm entry affects longer-run household investments and forward-looking decision-making. In the census, household investment in children's education provides one main measure, examined in Table 5. Pre-trends are small and statistically indistinguishable from zero throughout. The effects are concentrated at the secondary level. Primary enrolment, already very high at baseline (0.82 for girls and 0.84 for boys), is unchanged. In contrast, secondary enrolment rises sharply: among girls aged 14-18 it increases by 6.4 percentage points in the first treated period, a 39% increase from a base of 0.166. Boys' secondary enrolment rises by 4.2 points from a base of 0.157. For girls, the effect remains large one period later, at 8 percentage points, while the corresponding estimate for boys is 2.9 points.

Because enrolment only records whether a child is in school at the census date, we also test years of schooling, a stock measure of educational attainment for those aged 14-18 independent of current attendance. Girls report 0.21 additional years of schooling (from a base of 5.68 years), whereas boys' attainment does not significantly increase.

[38]Conditioning on the predictors of flower farm location reduces the lighting pre-trend without eliminating it (see Figure A10a). However, the networked structure of grid electrification implies that the cost of electrifying locations 'on the way' to higher electrification priorities is low. Our flower farm locations are likely to benefit from such externalities given the clustering of flower farms in the highlands.

Several mechanisms are consistent with these educational gains. Higher and more stable household income, in line with the labour market and wealth results, may ease the financial barriers to secondary schooling. The arrival of female-biased, stable formal-sector wage employment may also raise the expected returns to schooling; on its own, this channel would directly affect girls, whose future labour market potential the shock reshapes most. The positive effects on boys' enrolment point to broader forces, including an income effect or direct investment by the farms themselves. Industry reports document contributions to local schools in flower-growing areas (Kenya Flower Council, 2024), which would raise enrolment through supply rather than demand alone. Increases in the returns to education are not solely female-biased, either: men's wage employment also rises, potentially increasing boys' expected returns to education.

Table 5: Educational enrolment and attainment on flower farm entry

	(1) Primary enrol. (girls 6–13)	(2) Primary enrol. (boys 6–13)	(3) Sec. enrolment (girls 14–18)	(4) Sec. enrolment (boys 14–18)	(5) Years of schooling (girls 14–18)	(6) Years of schooling (boys 14–18)
First census after entry	-0.000	0.007	0.064***	0.042***	0.205**	0.110
	[0.013]	[0.010]	[0.015]	[0.014]	[0.091]	[0.076]
Second census after entry	-0.027	0.015	0.080***	0.029*	0.067	0.138
	[0.019]	[0.016]	[0.018]	[0.017]	[0.124]	[0.131]
Decade before entry (placebo)	-0.020	0.021	-0.021	0.002	0.070	0.038
	[0.028]	[0.018]	[0.031]	[0.026]	[0.171]	[0.132]
Never-treated mean (1999)	0.820	0.835	0.166	0.157	5.683	5.483
Early-treated sublocations	84	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55	55
Total sublocations	1,612	1,612	1,612	1,612	1,612	1,612
Observations	4,749	4,734	4,806	4,807	4,805	4,807

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates on a balanced panel of sublocations observed in the 1999, 2009 and 2019 censuses. Treatment definition and specifications as in Table 1. Household-level, averaged to the sublocation. Cols. (1)-(2) primary enrolment: household-level share of girls / boys aged 6–13 currently enrolled in school. Cols. (3)-(4) secondary enrolment: share of girls / boys aged 14–18 currently enrolled in secondary school. Cols. (5)-(6) average completed years of schooling of girls / boys aged 14–18. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The spatial-gradient plots show that the schooling effects are concentrated close to the farms (Figure A11). Effects are largest within 2-5km and become smaller by 10km.[39] This is qualitatively similar to the female wage-labour gradient documented above: in both cases the response is concentrated in near-farm sublocations rather than spread evenly across space.[40] Irrespective of the exact mechanism, educational attainment increases significantly, with educational choices of future workers much more aligned with urban labour market opportunities than traditional rural ones.

[39]Results are sensitive to the conditioning set: at 5km, estimates fall noticeably when conditioning on infrastructure, climate and water, or the full suitability score, although they remain positive and confidence intervals overlap across specifications. However, the spatial pattern remains unchanged.

[40]In the absence of spatially disaggregated data on beliefs about the returns to education, we cannot conclude from this spatial signature alone that supply-side or income effects dominate.

5.5 Family formation

The emergence of towns around flower farms provides two reasons to investigate family formation. First, the initial labour demand shock predominantly creates wage jobs for women, potentially affecting marriage and fertility decisions. Second, demographic choices in Kenya differ markedly between urban and rural households, raising the question of whether urbanisation itself changes family formation. If settlements become urbanised and households live in more urban conditions, do their choices also become more akin to those of urban households?

To address these questions, we examine whether the arrival of flower farms altered fertility choices and marriage patterns in treated areas. Table 6 reports findings separately for women at earlier (15-29) and later (30-44) stages of their reproductive lives, following standard convention in the literature. Panel A shows that any fertility effects are concentrated among younger women: following flower-farm entry, women aged 15-29 have 0.08 fewer children on average (column 1), are 2.3 percentage points less likely to have given birth in the previous three years (column 2), and 1.8 points less likely to have ever given birth (column 3). Relative to the never-treated means, these represent declines of around 6, 5, and 3% respectively. Figure A9 further shows that the decline in children ever born is concentrated within 2-5km of flower farms and fades by 10km, broadly mirroring spatial employment and education patterns. In contrast, the corresponding fertility estimates for women aged 30-44 are small and insignificant.[41]

Our findings are consistent with young women delaying family formation. A natural explanation is that flower-farm entry raises the opportunity cost of early childbearing. This may operate directly, by increasing the returns to young women's market work (e.g. Schultz, 1985), but also through the educational gains documented in Table 5. Increased female secondary school attainment may itself raise the returns to delaying marriage and childbearing (Heath and Jayachandran, 2018; Duflo et al., 2026). In fact, these channels are likely related: evidence from India and Bangladesh shows that improved employment opportunities for young women can increase educational attainment as well as delay family formation (Jensen, 2012; Heath and Mobarak, 2015). Our estimates may therefore reflect higher returns to women's work outside the family farm and business, as well as greater educational investment. These mechanisms are difficult to disentangle without additional exogenous variation and data on educational demand and supply.

[41]One caveat is that the fertility outcomes exhibit a positive pre-trend in our score-conditioned specification. Figure A9 shows that this pre-trend becomes more positive as we condition on more predictors of flower farm location. However, this pre-trend biases our finding of a decline in fertility following flower farm entry downward. In addition, Appendix A.IX.3 tests robustness of our results to violations of the parallel trends assumption following Rambachan and Roth (2023). The effect on children ever born remains significantly negative for violations of up to 85% of the observed pre-trend.

Table 6: Fertility and marriage on flower farm entry

	(1) Num. children ever born	(2) Gave birth (last 3 years)	(3) Ever given birth	(4) Currently married	(5) Ever divorced
	Panel A: Women aged 15–29				
First census after entry	-0.079***	-0.023**	-0.018*	-0.011	0.004
	[0.031]	[0.009]	[0.010]	[0.011]	[0.002]
Second census after entry	-0.006	0.007	-0.003	0.022	0.005
	[0.036]	[0.012]	[0.012]	[0.014]	[0.003]
Decade before entry (placebo)	0.098*	0.011	0.011	0.038**	0.001
	[0.050]	[0.016]	[0.016]	[0.016]	[0.004]
Never-treated mean (1999)	1.351	0.458	0.551	0.468	0.019
Early-treated sublocations	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55
Total sublocations	1,610	1,610	1,610	1,610	1,610
Observations	4,818	4,818	4,818	4,818	4,818
	Panel B: Women aged 30–44				
First census after entry	-0.025	-0.000	0.005	-0.003	0.014***
	[0.065]	[0.011]	[0.005]	[0.011]	[0.005]
Second census after entry	0.000	0.009	0.009	0.001	0.035***
	[0.094]	[0.015]	[0.007]	[0.012]	[0.008]
Decade before entry (placebo)	0.108	0.021	-0.013*	0.025	-0.005
	[0.113]	[0.019]	[0.007]	[0.019]	[0.012]
Never-treated mean (1999)	5.374	0.454	0.964	0.815	0.044
Early-treated sublocations	84	84	84	84	84
Late-treated sublocations	55	55	55	55	55
Total sublocations	1,609	1,609	1,609	1,609	1,609
Observations	4,815	4,815	4,815	4,815	4,815

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates on a balanced panel of sublocations observed in the 1999, 2009 and 2019 censuses. Treatment definition and specifications as in Table 1. Women of childbearing age (15–44); each panel balanced within its age group across the three census years. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The marriage results also fit such an interpretation. Young women become less likely to be currently married following flower-farm entry. Although merely a suggestive result, the negative, insignificant point estimate (column 4) contrasts notably with a significantly positive pre-trend in marriage rates.

Finally, Panel B of Table 6 points to a large increase in the probability that women aged 30-44 have ever divorced their partner. One obvious interpretation is that access to independent earnings from employment strengthens women's bargaining position and outside options within marriage. This interpretation is consistent with a broader literature linking women's economic opportunities to changes in household decision-making and marital stability. For example, increases in US women's economic independence have been associated with higher divorce rates (Stevenson and Wolfers, 2007), while more recent work shows that labour market shocks favouring women can reshape patterns of marriage and divorce (Autor et al., 2019). Divorce remains relatively rare, though: the 1.4 percentage point increase is relative to a never-treated mean of 4.4%.

5.6 Robustness

Our main results described above are highly robust to various definitions of the comparison group, the conditioning on floriculture suitability and different dynamic difference-and-differences estimation approaches. Figure 4 provided an overview of spatial gradients across different extents of the flower farm catchment area, separately by comparison group. Moving from the unconditional estimates through each of the three covariate blocks and, finally, to the full suitability score barely changes our main results.[42] Our Callaway and Sant'Anna (2021) estimates are also robust to various changes to our baseline specification, including restricting to only never-treated sublocations, dropping Nairobi district, or replacing the suitability score with its twelve individual component variables. We report first-census effects and placebos for our main outcomes reported above in Appendix A.IX.1.[43] Reassuringly, we can detect only marginal changes in point estimates. For example, the first-census-after-entry effect on female wage employment is 0.012 in our baseline and ranges from 0.009 to 0.013 across these alternatives, with almost completely overlapping confidence intervals (see Figure A12). The population results feature only weakly more variation in point estimates: the implied increase in the working-age population ranges from roughly 660 to 1,140 people (see Figure A13), but confidence intervals always overlap and the estimates remain significantly positive throughout. The increase in the urban share ranges from 4.7 to 5.4 percentage points, while the increase in private renting ranges from 2.5 to 3.6 points (see Figure A14). Housing and wealth outcomes are similarly stable across specifications (see Figure A16). Secondary enrolment among girls rises by 3.7 to 6.3 percentage points (see Figure A17), and the decline in children ever born among younger women ranges from 0.08 to 0.11 (see Figure A15).

The placebo estimates also barely change across specifications and generally remain close to zero. Wherever pre-trends are present, however, they persist across specifications, including for some fertility and marriage outcomes. We return to these below, using the methodology proposed by Rambachan and Roth (2023) to assess the sensitivity of the post-treatment estimates to violations of the parallel trends assumption.

One mild exception is the specification directly adjusting for all twelve components of the score instead of the score itself. With only 139 treated sublocations, the direct adjustment on twelve separate covariates is very restrictive and concentrates comparison weights on relatively few places. Estimates are qualitatively unchanged, but are the one specification that differs quantitatively, with a tendency to attenuate treatment effects toward zero, for example for girls' secondary school enrolment.

[42] Appendix A.VIII provides all main outcomes' spatial gradients across all comparison groups.
[43] Full regression tables are reported in the Online Supplement.

We also check robustness of our results to the restriction on common-support for the full suitability score in Appendix A.IX.2. We re-estimate female wage employment and population under a range of thresholds, from no trimming of the least suitable sublocations to trimming up to the tenth percentile of the treated sublocations' suitability score. Point estimates barely change. If anything, the first-census effect on female wage employment increases as the threshold becomes tighter, while the population effect decreases somewhat. Our trimming choice of the minimum suitability score of the treated sublocations sits safely in the middle of the various trimming assumptions.

To take potential violations of the conditional parallel trends assumption seriously, we assess the sensitivity of the five outcomes with significant ($p < 0.10$) placebo estimates using the approach proposed by Rambachan and Roth (2023) in Appendix A.IX.3. We restrict post-treatment violations of parallel trends to the direction of each outcome's observed pre-trend. For the five problematic outcomes, results are reassuring: the decline in children ever born among women 15-29 remains bounded away from zero as long as violations are up to half the size of the observed pre-trend, and only marginally includes zero when they are equally large. The electric-lighting effect is similarly robust, while the piped-water effect remains positive throughout. The other two outcomes (currently married and ever given birth) have post-treatment estimates that are themselves indistinguishable from zero, so the significant placebos do not affect our interpretation of these results. Taken together, the sensitivity analysis supports our cautious interpretation of the significant effects on fertility and housing investments.

Finally, to show robustness of our main results to different estimation approaches, we re-estimate the employment and population results using the event-study estimator of de Chaisemartin and D'Haultfœuille (2024) in Appendix A.IX.4. This estimator enables us to study the dynamic difference-in-differences effect for continuous definitions of treatment such as greenhouse area (intensive margin) in addition to the effect of greenhouse indicator (extensive margin). Results for the latter are essentially unchanged, with first-census estimates being close to their Callaway and Sant'Anna (2021) counterparts.[44] Regarding the former, the intensive margin estimates for greenhouse area confirm the documented pattern for the extensive margin: an additional 10 hectares of nearby greenhouse area increases women's wage employment by 0.3 percentage points and the local working-age population by around 190 people in the first census after entry. We interpret these results as showing that both the extensive and intensive margins of flower farms matter for town formation.[45]

[44]Note that the second-census estimates are not directly comparable in levels because de Chaisemartin and D'Haultfœuille (2024) divide by cumulative treatment since entry. Their second-census coefficient is therefore roughly half the corresponding Callaway and Sant'Anna (2021) estimate.

[45]We leave a more thorough investigation of potential non-linearities in town formation as a function of the size of the labour demand shock for future work.

6 Mechanisms

How does an agro-industrial labour demand shock generate a town? Once workers agglomerate and draw in further population to form a small town, can it sustain itself as a centre of economic activity even without (the growth of) the initial employer? To motivate questions around persistence, it is important to remember that in our context of flower farms in the Kenyan highlands, these factory-like employers provide more than just jobs: beyond the income effect of flower farm employment and the attraction of migrant workers, flower farms also make or help attract broader investments that can improve the productivity and amenity value of their host locations. This includes the provision of public services including schools, health facilities and nurseries, as well as public infrastructure such as bus services, wells, electricity and minor road upgrades. Taken together, flower farm arrival represents a local 'big push' investment that could transform a location in a persistent manner (Murphy et al., 1989).[46]

In the following, we provide three pieces of suggestive evidence to address questions on how a town may form, and whether it can persist. First, we document that flower-farm entry generates employment and structural transformation beyond the farms themselves, as local demand draws workers into other occupations. Second, we show that flower farm workers and other residents make investments that can sustain income and consumption gains into the future, further reinforcing the attractiveness of the town. Finally, we study potential persistence more directly by investigating instances of flower farm closure. Taken together, these pieces of evidence paint a coherent picture that the initial employment shock sets in motion a long-run process of local development that becomes increasingly self-sustaining.

6.1 Occupational diversification

The urbanisation and population growth documented above suggest that the impacts of rural proto-industrialisation extend beyond the direct creation of wage jobs on flower farms. By stimulating demand for complementary goods and services, new flower farm jobs may generate indirect employment opportunities and accelerate the reallocation of workers away from agriculture. To explore whether new towns formed after flower farm entry exhibit diversification of their occupational structure, we exploit the rich occupational information available in the 2019 census microdata.

Table 7 compares the sectoral composition of employment in 2019 across sublocations that did and did not receive a flower farm at any time between 2000 and 2019,

[46] A natural concern with 'big push' interventions is that general equilibrium effects may undermine SUTVA. However, a given flower farm can attract hundreds of individual migrants from dispersed settlements across Kenya. Most migrants will have previously been self-employed on their own fields, and only a very small number from any given settlement would leave. Hence, any adverse effect of departing migrant workers on their origin locations is very limited.

controlling for baseline urbanisation in 1999. Despite flower farms providing agricultural employment, treated sublocations exhibit substantially lower employment shares in agriculture for both women and men: among the employed, the share of women (men) working in agriculture is 7.7 (9.6) percentage points lower in communities located within 5km of a flower farm. Compared to mean values of the never-treated, these effects represent sizeable relative declines of approximately 9 and 14%, respectively.

Table 7: Sectoral employment shares on greenhouse presence within 5km at endline (2019).

	(1) Agriculture	(2) Manufacturing & construction	(3) Transport & storage	(4) Wholesale & retail	(5) Accommodation & food	(6) Health & education	(7) Other
				Panel A: Women			
Flower farm within 5km	-0.077***	0.004*	0.000	0.020***	0.010***	0.011***	0.032***
	[0.018]	[0.002]	[0.000]	[0.006]	[0.002]	[0.003]	[0.009]
Never-treated mean	0.815	0.014	0.001	0.058	0.019	0.032	0.061
R-squared	0.499	0.113	0.078	0.298	0.267	0.128	0.480
				Panel B: Men			
Flower farm within 5km	-0.096***	0.025***	0.020***	0.015***	0.006***	-0.001	0.031***
	[0.017]	[0.006]	[0.003]	[0.005]	[0.002]	[0.002]	[0.007]
Never-treated mean	0.679	0.074	0.055	0.052	0.013	0.031	0.096
R-squared	0.338	0.092	0.090	0.295	0.104	0.031	0.241
Treated sublocations	139	139	139	139	139	139	139
Sublocations	5,377	5,377	5,377	5,377	5,377	5,377	5,377

Notes: OLS at the sublocation level (one observation per balanced-panel sublocation): The dependent variable in each column is the share of employed residents aged 15–50 (2019 census) whose main job falls in that sector group; the sample is restricted to those with a reported occupation, so the seven shares sum to one. Sector groups follow ISIC; column (2) combines manufacturing with construction, column (6) health with education, and column (7) pools all remaining sectors (mining, utilities, finance, real estate, professional services, public administration, and other services). See Appendix A.VI for more details on sector classification. Sublocations with a greenhouse within 5km already by 1999 are excluded; among the remainder the greenhouse indicator equals 1 for sublocations first treated between 2000 and 2019 (never-treated omitted). All regressions control for the sublocation's 1999 baseline urban share (the share of residents aged 15–50 classified urban in the 1999 census; coefficient not reported). Never-treated mean is the mean sector share across never-treated sublocations. Robust standard errors in brackets. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

For women, the lower share of agricultural employment is mirrored by higher employment shares across a broad range of market-oriented activities. The largest increases are in wholesale and retail trade, accommodation and food services, health and education, and a residual category combining smaller service occupations. For men, the shift out of agriculture is accompanied by higher employment shares in manufacturing and construction, transport and storage, wholesale and retail trade, accommodation and food services, and other sectors. Whereas the increase in transport and storage is potentially linked to the logistical needs of export-oriented, perishable cut-flower production, the gains in manufacturing, construction and wholesale/retail appear to document structural transformation beyond the value chain of flower farms.

These estimates are cross-sectional only, and should not be interpreted causally. Nonetheless, they indicate broader occupational diversification in flower farm towns.

That towns around flower farms support a more diverse economy can be interpreted as suggestive of a concentration of economic activity that may become self-sustaining. However, diversification alone is not sufficient evidence of persistence: supply-chain linkages and demand from relatively affluent flower-farm workers could generate growth in other sectors that remains dependent on the farms and may therefore disappear if they close. A stronger test of persistence is whether such economic activity survives the loss of the town's initial comparative advantage (Bleakley and Lin, 2012).

Since Kenya's flower boom is ongoing, we cannot test this directly for most flower-farm towns. We instead provide two further pieces of suggestive evidence. Section 6.2 examines whether residents make forward-looking investments that may help sustain local economic activity beyond the initial employment shock, while Section 6.3 provides a case study of two large flower farm closures, documenting changes in a nearby town that loses the majority of its original source of flower farm employment.

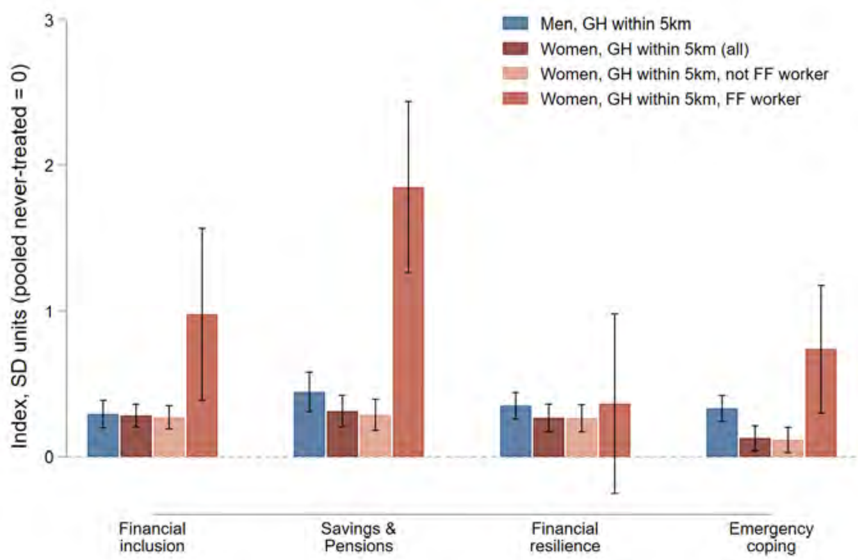
6.2 Forward-looking investments

One micro-foundation of future persistence is forward-looking investments that tie inhabitants more closely to their location. We exploit unusually rich individual-level survey data obtained by the Kenya Bureau of National Statistics on the financial lives of a nationally representative sample of 20,871 individuals. We use the 2024 FinAccess survey to study individual-level measures of investment and financial access to test whether residents near flower farms are more financially included, save more for the future, and invest more in less mobile assets, e.g., land, housing, and businesses. Since the data are a single geo-identifiable cross-section, results are interpreted as suggestive.

Figure 6 summarises four dimensions of financial wellbeing: financial inclusion; savings and pensions; financial resilience, based on reported financial hardship; and emergency coping, based on how households respond to unexpected expenses.[47] Curiously, the endline survey also contains instances of workers who report working in a flower farm. This allows us to compare not only individuals living near flower farms to those far from them, but also to separate women who report working for a flower farm in treatment locations from other working women. Each figure bar shows the level of the index for that respective group relative to the never-treated benchmark, conditional on baseline urbanisation, age, and marital status.

[47] Each dimension is represented by an index constructed from several individual-level questions. Appendix A.XI provides details on index construction; Table A11 reports the question-level estimates.

Figure 6: Financial indices by greenhouse proximity and flower farm employment



Notes: FinAccess 2024, respondent level. Each bar is the group's never-treated mean plus the coefficient on a greenhouse-within-5 km indicator from a separate regression for that group, controlling for 1999 sublocation urban share, age and marital status. Indices are in never-treated standard deviations. Whiskers: 95% confidence intervals, standard errors clustered at the sublocation level. *Treated* includes all ever-treated sublocations. Table A12 reports full underlying estimates, distinguishing flower-farm workers from other residents separately for women and men.

Relative to respondents in never-treated sublocations, women employed on flower farms score roughly 1.0 standard deviations (SD) higher on financial inclusion and 1.9 SD higher on savings and pensions. It is important to note that there are only 15 female flower farm workers in the 2024 FinAccess sample, spread across 11 sublocations, so differences are noisy, but large.[48] The relative stability of flower-farm employment may help explain these large differences.[49] Importantly, the higher scores are not confined to flower-farm workers. Women living within 5km of a flower farm who work outside of the sector score around 0.3 SD higher on both financial inclusion and savings and pensions. Men living near farms show similarly sized differences.

Table 8 investigates forward-looking investments in the form of less mobile capital. Given the low frequency of these investments, we pool all individuals at the sublocation-level to increase statistical power and condition on baseline urbanisation status to reduce noise. We find that residents near flower farms are 0.8 percentage points more likely to have bought land in the previous year, against a never-treated mean of 1.8%. Spending on housing materials rises by 1.3 points, from a 5.1% mean, while purchases

[48]Large effects on wage-employed women's financial situation are in line with evidence that women's access to their own earnings can strengthen their economic position and household bargaining power (Field et al., 2021; Jayachandran and Voena, 2026).

[49]The size of the savings differences is noteworthy given the mixed evidence on industrial jobs elsewhere in Africa. Experimental studies in Ethiopia find that factory employment can raise savings (Abebe et al., 2025; Kotsadam et al., 2026), but also document high worker turnover and, in some settings, limited longer-run gains (Blattman et al., 2022). One possible difference is persistence in the jobs themselves: sustained employment gives workers more scope to turn wages into savings and assets.

of business equipment more than double, from 1.2% to 2.5%. In contrast, livestock purchases fall by 2.5 points, from a base of 12.8%. These suggestive results are consistent with greater investment in immobile assets that tie individuals and households to the town, in line with the causal improvements in housing quality documented in Section 5.3. The shift from livestock towards business assets also independently confirms the broader move out of agriculture documented in Section 6.1 above.

Table 8: Major purchases in the past 12 months, by greenhouse proximity

	(1) Land	(2) New house	(3) House materials	(4) Business equipment	(5) Farm equipment	(6) Livestock
Greenhouse within 5km (by 2019)	0.008*	0.003	0.013*	0.013***	0.006	-0.025**
	[0.004]	[0.003]	[0.007]	[0.005]	[0.005]	[0.010]
Observations	1443	1443	1443	1443	1443	1443
Never-treated mean	0.018	0.008	0.051	0.012	0.021	0.128
Treated sublocs	117	117	117	117	117	117
Never-treated sublocs	1,336	1,336	1,336	1,336	1,336	1,336

Notes: Cross-sectional OLS, FinAccess 2024. Women and men pooled. 5km greenhouse buffer; sublocation is the unit of analysis (outcomes are sublocation means over respondents of both sexes). Outcomes: whether the respondent made each major purchase in the past 12 months. Regressor: binary treated indicator (never-treated omitted). The only control included is urban (1999 sublocation census share); the sublocation-mean age and marital-status controls of the main table are omitted, since in-migration plausibly shifts both. Robust standard errors. * p<0.10, ** p<0.05, *** p<0.01.

6.3 Flower farm closures

Naivasha, Kenya's flower-growing capital, provides an example of what a mature flower town could ideally become. Its rapid growth into a secondary city has been closely tied to the expansion of export floriculture around Lake Naivasha (Kuiper and Greiner, 2021). However, whether the small towns that emerge in response to the arrival of flower farms can eventually sustain themselves is crucial for our interpretation of the long-run development effects of their labour demand shocks. A natural test of the persistence of these towns is to study sudden job loss from flower farm closures.

Our greenhouse data identifies eleven potential farm closures, of which two qualify as case studies.[50] The two large neighbouring operations, whose closure we can independently verify via press reports, are both near Kericho town in Kericho County.[51]

Therefore, we use Kericho town as a case study of what happens when an established flower-growing area experiences a large negative employment shock. As shown in Appendix A.XII, two out of three flower farms surrounding the town closed in 2019, with visible greenhouse area collapsing from 130 to 20 hectares in the four years until

[50] Of the initial eleven cases, three are false positives, with greenhouses disappearing temporarily before reappearing; three others involve very small operations that closed within a few years of opening; another three were sizeable farms, but located alongside other major flower farms that continued operating, leaving us with two confirmed cases of closure to investigate further (see Online Supplement).

[51] Appendix A.XII provides imagery for the two case study farm closures we focus on.

2023. Both closed flower farms were operated by the same owner, Finlays.[52] Announced in October 2019, the large farms were quickly liquidated. About 1,700-2,000 workers lost their jobs in the process. One farm disappeared completely, appearing fully reforested on 2024 satellite imagery, whereas the second site attracted a botanical farm to use a part of its former land in 2024. Lacking post-2019 census data, we employ maximum nighttime luminosity, which is particularly sensitive to peak light emissions, as a high frequency proxy for changes in economic activity in Kericho around the closures.

We find a stark drop in nighttime luminosity in 2021 that tracks the collapse in greenhouse area around the town with a one-to-two-year lag.[53] However, one year after the nightlights decrease, nighttime luminosity increases again in 2022 and reverts back to higher maximum nighttime luminosity than before the closures by 2023. On satellite imagery of the area, no noticeable change in spatial extent or agglomeration of the town is visible five years after the adverse employment shock.

We cautiously interpret this finding as highlighting the potential of agro-industrial labour shocks to generate self-sustaining towns in the sense of Bleakley and Lin (2012). It is important to note, however, that a case study of two flower farm closures in one town does not identify any average effect of farm closure. Nonetheless, it provides one illustration of the potential persistence of flower towns, as the local economy appears to have survived the loss of two major employers. This is consistent with a more self-sustaining centre of economic activity, which reacts somewhat asymmetrically to the arrival and departure of a large wage employer, with the settlement becoming less dependent on the shock that transformed it into a town in the first place.

7 Conclusion

What happens to rural economies when stable wage jobs arrive at scale? We study the expansion of Kenya's cut-flower industry since 2000, which brought large numbers of predominantly female wage jobs to the Kenyan highlands, and trace how initially rural settlements transform over the following two decades. Exploiting the agroclimatic, geological and economic requirements of cut-flower production, we employ a matched difference-in-differences design to compare areas around newly established flower farms with similarly suitable locations where no flower farm entered.

We find that the arrival of these wage jobs sets in motion a broader process of urbanisation and town formation. Population grows substantially around the farms, driven in large part by in-migration, and continues to increase over time. The structure

[52] The reasons for the closures, according to a Finlays general manager, were a weakening exchange rate, extreme weather events and high labour cost. In addition, at least one source mentions recent strikes at both farms which highlighted intractable tension between management and the workforce.

[53] This drop remains when we residualise nighttime lights on year fixed effects, defusing concerns that the Kericho effect could be driven by Covid-19, or wider issues with electricity outages.

of the local economy changes, with women supplying labour to the market and moving out of self-employment. Settlements become more likely to be officially classified as urban. Rental markets expand, housing quality and utility access improve, and household wealth rises. Families invest more in education, and younger women appear to delay family formation. Taken together, these results show that the agro-industrial labour demand shock caused by flower farm entry in rural Kenya created new centres of economic activity: blossoming towns.

We also provide suggestive evidence on whether these newly formed agglomerations can become self-sustaining towns. Although the initial employment shock comes from agricultural employment in floriculture, employment patterns at endline are consistent with occupational diversification away from agriculture. We also find greater savings and pension participation, alongside durable investments in land, housing, and businesses. A case study of two major farm closures provides an example of town persistence beyond the flower farms: economic activity initially falls, but subsequently recovers, while the town itself shows no visible signs of contraction five years later. Overall, these patterns suggest that the changes set in motion by flower-farm entry may persist as local economies become more diversified and less dependent on the initial source of employment.

Historical comparisons with proto-urbanisation in Europe highlight that the agro-industrial jobs that flower farms bring to rural areas are not unique. However, they appear promising as engines of transformation and urbanisation in contemporary sub-Saharan Africa: whereas non-agricultural wage employment remains heavily concentrated in a few large cities, flower farms provide wage jobs in rural locations that satisfy their input requirements. Finally, their labour-intensive, industrial organisation requires large numbers of workers, including migrants, whose recruitment is supported by the private provision of infrastructure and services, following historical precedent in the European textile industry.

Why do rural, agrarian settlements like the ones we study in the Kenyan highlands not agglomerate into towns in the absence of flower farms? Bryan et al. (2025) document recent increases in population density in many rural parts of Africa, which we interpret as indicative of one of two phenomena: either a high latent potential in rural areas for the concentration of population into towns, or the early stages of a new era of endogenous agglomeration. The former would imply that attracting rural employers can have large externalities by helping to coordinate on location. The latter would imply that flower farms represent one example of a broader trend across the continent. In either case, understanding the determinants of rural agglomeration and identifying the potential frictions impeding it appear promising avenues for future research.

References

Abebe, Girum, Niklas Buehren, and Markus Goldstein, "Do Factory Jobs Improve Welfare? Experimental Evidence from Ethiopia," *The World Bank Economic Review*, 2025, *39* (1), 124–142.

Acampora, Michelle, Lorenzo Casaburi, and Jack Willis, "Land Rental Markets: Experimental Evidence from Kenya," *American Economic Review*, March 2025, *115* (3), 727–71.

Ahlfeldt, Gabriel M. and Elisabetta Pietrostefani, "The economic effects of density: A synthesis," *Journal of Urban Economics*, 2019, *111*, 93–107.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press, 2009.

Arvis, Jean-François, John Arnold, Heinrich Bofinger, and Michael Webber, "Air Freight: A Market Study with Implications for Landlocked Countries," *World Bank TP*, 2009, (26).

Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review*, 2020, *110* (3), 797–823.

Ashmore, Owen, *The Industrial Archaeology of North-West England*, Vol. 29 of *Remains, Historical and Literary, Connected with the Palatine Counties of Lancaster and Chester*, Manchester: Manchester University Press, 1982.

Autor, David H., David Dorn, and Gordon H. Hanson, "When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Young Men," *American Economic Review: Insights*, 2019, *1* (2), 161–178.

Bandiera, Oriana, Ahmed Elsayed, Andrea Smurra, and Céline Zipfel, "Young Adults and Labor Markets in Africa," *Journal of Economic Perspectives*, February 2022, *36* (1), 81–100.

Blattman, Christopher, Stefan Dercon, and Simon Franklin, "Impacts of Industrial and Entrepreneurial Jobs on Youth: 5-Year Experimental Evidence on Factory Job Offers and Cash Grants in Ethiopia," *Journal of Development Economics*, 2022, *156*, 102807.

Bleakley, Hoyt and Jeffrey Lin, "Portage and Path Dependence," *The Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

Bryan, Gharad and Melanie Morten, "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

__ , Kyra Frye, and Melanie Morten, "Spatial economics for low- and middle-income countries," in Dave Donaldson and Steve Redding, eds., *Handbook of Regional and Urban Economics*, Vol. 6, Amsterdam: Elsevier North Holland, 2025, chapter 10, pp. 653–730.

Button, Kenneth, "Air Freight Transport and Economic Development," *Journal of Air Transport Management*, 2020.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230. Themed Issue: Treatment Effect 1.

Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio, "What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States," *Journal of Urban Economics*, 2017, *98*, 17–49.

Chen, Xiuxiu, Zeyu Wang, Feng Zhang, Guoqiang Shen, and Qiuxiao Chen, "A global annual simulated VIIRS nighttime light dataset from 1992 to 2023," 2024, *11* (1), 1380.

Chiovelli, Giorgio, Stelios Michalopoulos, Elias Papaioannou, and Tanner Regan, "Illuminating the Global South," *The Economic Journal*, 2026, p. ueaf134.

Ciccone, Antonio and Robert E. Hall, "Productivity and the Density of Economic Activity," *The American Economic Review*, 1996, *86* (1), 54–70.

Davis, Donald R. and David E. Weinstein, "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, December 2002, *92* (5), 1269–1289.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *The Review of Economics and Statistics*, 2024, pp. 1–45.

Dehejia, Rajeev H. and Sadek Wahba, "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *J. Am. Stat. Assoc.*, 1999, *94* (448), 1053–1062.

__ **and** __ , "Propensity Score-Matching Methods for Nonexperimental Causal Studies," *The Review of Economics and Statistics*, 2002, *84* (1), 151–161.

Dinkelman, Taryn and L. Rachel Ngai, "Time Use and Gender in Africa in Times of Structural Transformation," *Journal of Economic Perspectives*, February 2022, *36* (1), 57–80.

Dolan, Catherine, Maggie Opondo, and Sally Smith, "Gender, Rights and Participation in the Kenya Cut Flower Industry," *SSR Project No. R8077 2002-4*, 2002.

Duflo, Esther, Pascaline Dupas, Elizabeth Spelke, and Mark P. Walsh, "Intergenerational Impacts of Secondary Education: Experimental Evidence from Ghana," *American Economic Review*, 2026. Forthcoming.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker, "General equilibrium effects of cash transfers: Experimental evidence from Kenya," *Econometrica*, 2022, *90* (6), 2603–2643.

Fick, Stephen E. and Robert J. Hijmans, "WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas," *International Journal of Climatology*, 2017, *37* (12), 4302–4315.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore, "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms," *American Economic Review*, 2021, *111* (7), 2342–2375.

Figueiredo Walter, Torsten and Niclas Moneke, "When Does Electrification Work? Evidence from Sub-Saharan Africa," 2026. University of Oxford mimeo.

Food and Agriculture Organization, "Water Bodies: Kenya, 2000 (Africover)," `https://purl.stanford.edu/gt901vg7062` 2000. Distributed by Stanford Digital Repository.

Glaeser, Edward L. and Joshua D. Gottlieb, "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the US," *Journal of Economic Literature*, 2009, *47* (4), 983–1028.

Gollin, Douglas and Joseph P. Kaboski, "New views of structural transformation: insights from recent literature," *Oxford Development Studies*, 2023, *51* (4), 339–361.

__ **and Richard Rogerson**, "Productivity, transport costs and subsistence agriculture," *Journal of Development Economics*, 2014, *107*.

__ **, David Lagakos, and Michael Waugh**, "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*, 2014, *129* (2), 939–993.

__ **, Remi Jedwab, and Dietrich Vollrath**, "Urbanization with and without industrialization," *Journal of Economic Growth*, 2016, *21*, 35–70.

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

Harari, Mariaflavia, "Cities in Bad Shape: Urban Geometry in India," *American Economic Review*, August 2020, *110* (8), 2377–2421.

Heath, Rachel and A. Mushfiq Mobarak, "Manufacturing Growth and the Lives of Bangladeshi Women," *Journal of Development Economics*, 2015, *115*, 1 – 15.

__ **and Seema Jayachandran**, "The Causes and Consequences of Increased Female Education and Labour Force Participation in Developing Countries," in "The Oxford Handbook of Women and the Economy," Oxford University Press, 2018.

Heblich, Stephan, David Krisztián Nagy, Alex Trew, and Yanos Zylberberg, "The Death and Life of Great British Cities," May 2026. Unpublished manuscript.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd, "Matching as an Econometric Evaluation Estimator," *The Review of Economic Studies*, 1998, *65* (2), 261–294.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, April 2012, *102* (2), 994–1028.

__ **and Vivian Liu**, "Urban Land Markets and City Development: Sub-Saharan Africa," Technical Report 44, Department of Geography and Environment, LSE 2023.

— , **Francisco Libano-Monteiro, Martina Manara, Guy Michaels, and Tanner Regan**, "Economics of Greenfield Urban Planning," Working Papers 2025-007, The George Washington University, The Center for Economic Research Oct 2025.

— , **Tanner Regan, and Anthony J Venables**, "Building the City: From Slums to a Modern Metropolis," 2021, *88* (3), 1157–1192.

Hengl, Tomislav, Johan G. B. Leenaars, Keith D. Shepherd, Markus G. Walsh, Gerard B. M. Heuvelink, Tekalign Mamo, Helina Tilahun, Ezra Berkhout, Matthew Cooper, Eric Fegraus, Ichsani Wheeler, and Nketia A. Kwabena, "Soil Nutrient Maps of Sub-Saharan Africa: Assessment of Soil Nutrient Content at 250 m Spatial Resolution Using Machine Learning," *Nutrient Cycling in Agroecosystems*, 2017, *109* (1), 77–102.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi, "Growth and Structural Transformation. Chapter 6," *Handbook of Economic Growth*, 2014, *2*, 855–941.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," 2014, *129* (4), 1899–1946.

Hoskins, W. G., *The making of the English landscape*, London: Hodder and Stoughton, 1955.

IFC, "How Kenya's Flower Industry Is Growing More Sustainable," 2020. Accessed: 01/08/26.

Jaffee, Steven M., "How Private Enterprise Organized Agricultural Markets in Kenya," Policy Research Working Paper 823, World Bank, Washington, DC 1992.

Jayachandran, Seema and Alessandra Voena, "Women's Power in the Household," *Journal of Economic Literature*, 2026, *64* (2), 447–497.

Jensen, Robert, "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *The Quarterly Journal of Economics*, 03 2012, *127* (2), 753–792.

Kazimierczuk, AH, P Kamau, BK Kinuthia, and C Mukoko, "Never a rose without a prick: (Dutch) multinational companies and productive employment in the Kenyan flower sector," *African Studies Centre Working Paper Series*, 2018.

Kenya Flower Council, "The Kenyan Floriculture Sector 2024: A Review of Floriculture Sector in Kenya," Floriculture Sustainability Assessment Report, Naivasha, Kenya 2024.

Kline, Patrick and Enrico Moretti, "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kotsadam, Andreas, Janneke Pieters, Pauline Rossi, Espen Villanger, Céline Zipfel, and Tigabu Getahun, "Female Jobs and Fertility: Long-term Experimental Evidence from Ethiopian Factories," 2026. Working paper.

Krugman, Paul, "Increasing Returns and Economic Geography," *The Journal of Political Economy*, 1991, *99* (3).

— , *Development, geography, and economic theory* 1995.

Ksoll, Christopher, Rocco Macchiavello, and Ameet Morjaria, "Electoral Violence and Supply Chain Disruptions in Kenya's Floriculture Industry," 2023, *105* (6), 1335–1351.

Kuiper, Gerda, *Agro-industrial Labour in Kenya: Cut Flower Farms and Migrant Workers' Settlements*, Cham: Springer International Publishing, 2019.

— **and Clemens Greiner**, "Export horticulture and labour migration in Kenya: Translocality and transiency in a secondary town," *Geoforum*, 2021, *122*, 1–10.

Lagakos, David, "Urban-Rural Gaps in the Developing World: Does Internal Migration Offer Opportunities?," *Journal of Economic Perspectives*, August 2020, *34* (3), 174–92.

Macchiavello, Rocco and Ameet Morjaria, "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports," *American Economic Review*, 2015, *105* (9), 2911–45.

MacDonald, Alan M., Helen C. Bonsor, Brighid É. Ó Dochartaigh, and Richard G. Taylor, "Quantitative Maps of Groundwater Resources in Africa," *Environ. Res. Lett.*, 2012, *7* (2).

Marshall, Alfred, *Principles of Economics*, London: Macmillan and Co., 1890.

Meijer, Johan R., Mark A. J. Huijbregts, Kees C. G. J. Schotten, and Aafke M. Schipper, "Global Patterns of Current and Future Road Infrastructure," *Environ. Res. Lett.*, 2018, *13* (6).

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding, "Urbanization and Structural Transformation," *The Quarterly Journal of Economics*, 05 2012, *127* (2), 535–586.

Mitullah, Winnie, Paul Kamau, and Joshua M. Kivuva, "Employment Creation in Agriculture & Agro-Processing Sector in Kenya," *PASGR Working Paper 020*, 2017.

Moneke, Niclas, "Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia," 2023. University of Oxford mimeo.

Müller-Crepon, Carl, "State Reach and Development in Africa since the 1960s: New Data and Analysis," *Political Science Research and Methods*, 2023, *11* (3), 663–672.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "Industrialization and the Big Push," *Journal of Political Economy*, 1989, *97* (5), 1003–1026.

Nagy, Dávid Krisztián, "Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion," *The Review of Economic Studies*, 11 2023, *90* (6), 3238–3281.

Observatory of Economic Complexity, "Cut Flowers in Kenya," 2023.

OECD, "Africa's Urbanisation Dynamics 2022: The Economic Power of Africa's Cities," Technical Report, OECD Publishing, Paris 2022.

Princeton University Library, "Africa Road Network, 1990s (digitized from Michelin maps, 1:4,000,000)," `https://maps.princeton.edu` 2021.

Rambachan, Ashesh and Jonathan Roth, "A More Credible Approach to Parallel Trends," *Review of Economic Studies*, 2023, *90* (5), 2555–2591.

Rikken, Milco, "The Global Competitiveness of the Kenyan Flower Industry," Technical Paper, World Bank Group and Kenya Flower Council December 2011.

Rosenstein-Rodan, Paul, "Problems of Industrialization of Eastern and Southeastern Europe," *Economic Journal*, 1943, *53*, 202–211.

Sant'Anna, Pedro H. C. and Jun Zhao, "Doubly Robust Difference-in-Differences Estimators," *Journal of Econometrics*, 2020, *219* (1), 101–122.

Schultz, T. Paul, "Changing World Prices, Women's Wages, and the Fertility Transition: Sweden, 1860-1910," *Journal of Political Economy*, 1985, *93* (6), 1126–1154.

Shin, Yeong Chan, Je Yeon Yeon, and Wan Soon Kim, "Influence of Suboptimal Temperature on Flower Quality and Floral Organ Development in Spray-Type Cut Rose 'Pink Shine'," *Horticulturae*, 2023, *9* (8), 861.

Stevenson, Betsey and Justin Wolfers, "Marriage and Divorce: Changes and their Driving Forces," *Journal of Economic Perspectives*, 2007, *21* (2), 27–52.

Wadsworth, Alfred P. and Julia De Lacy Mann, *The Cotton Trade and Industrial Lancashire, 1600–1780* number 7. In 'Publications of the University of Manchester. Economic History Series.', Manchester: Manchester University Press, 1931.

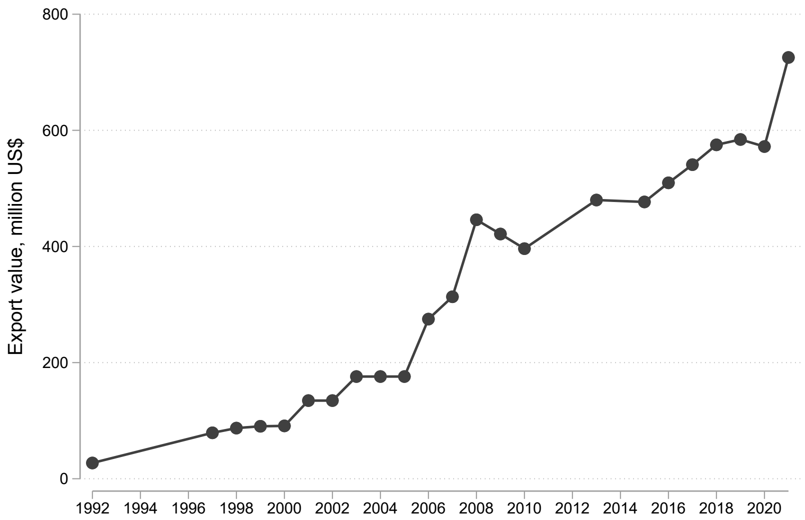
World Bank, "World Development Indicators," 2026. Series `NY.GDP.PCAP.PP.KD`, `NY.GDP.PCAP.KD` and `SP.URB.TOTL.IN.ZS`. Accessed 17 August 2026.

World Bank ESMAP and Solargis, "Global Solar Atlas 2.0," `https://globalsolaratlas.info` 2019.

Online Appendix

A.I Background

Figure A1: Export revenues from cut flower sector in Kenya: 1992-2021.



Notes: Annual value of Kenyan cut-flower exports (HS code 0603), 1992–2021, in millions of current US dollars. *Source:* UN Comtrade (2023).

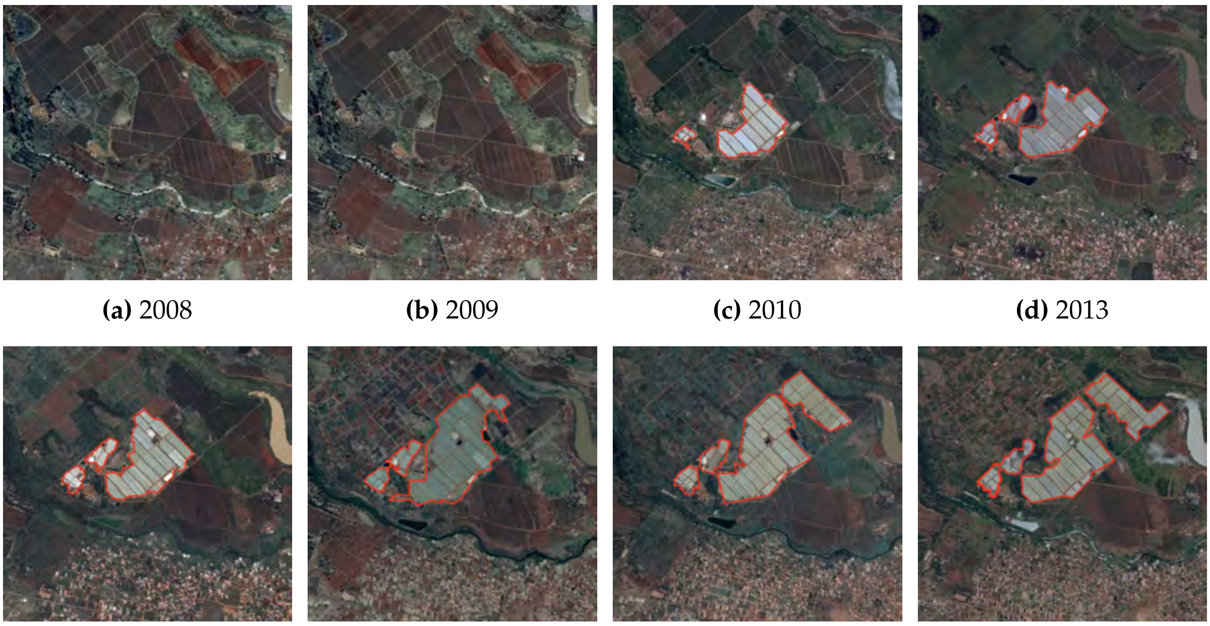
A.II Measuring greenhouse arrival and expansion

Figure A2 illustrates our detection and measurement procedure for a representative flower farm in the sample. For consistency, we use the same example as shown in Figure 1. Using the most recent available imagery for the part of Kenya that contains this (group of) greenhouses, in this case from 2022, the flower farm constituted by the group of greenhouses is clearly visible to the human eye. After noting down its centroid GPS coordinates, we search backwards through historic imagery to determine the time of entry. In this case, the sequence of Google Earth images shows no greenhouse structures in 2008 or 2009, a first appearance in 2010, and their subsequent expansion through 2022. The contours superimposed on each image represent our manual delineation of the greenhouse boundaries for each available image, from which we derive area covered by greenhouses per flower farm in a given year. Taken together, both form the basis of our measures of farm entry and intensive-margin growth. At the sublocation-level, we can add the number of (groups of) greenhouses as a second intensive margin measure of treatment intensity.

Two caveats deserve attention. First, in practice the year of entry is defined as the first year in which greenhouse structures become visible in historical satellite imagery. Although coverage is generally sufficiently rich to reconstruct the timing and expansion of flower farms, cloud cover and other imperfections occasionally create gaps in the image sequence. Consequently, the recorded year of entry may differ from the true year by one or two years due to gaps in historical satellite coverage. In the annual night-light event studies, this uncertainty introduces some classical measurement error in treatment timing. By contrast, the coarser temporal variation in the census makes the distinction between early- and late-entry cohorts relatively robust to such error.

Second, while greenhouses themselves are visible from space, the crops cultivated inside them are not. This raises the concern that our sample greenhouses may be growing products other than flowers. We address this concern by collecting information on the companies associated with the greenhouses using Google Maps, where possible. Where websites detailing company activities are not available, company names often reveal the likely products grown. Cut flowers, and especially roses, represent the dominant form of cultivation for this sample of cross-verifiable greenhouses. In addition, qualitative evidence with flower farm managers, anecdotal evidence by KFC experts, and our own field observation lead us to assume that all greenhouses in our sample are involved in the cultivation of flowers.

Figure A2: Measuring greenhouse entry and expansion over time.



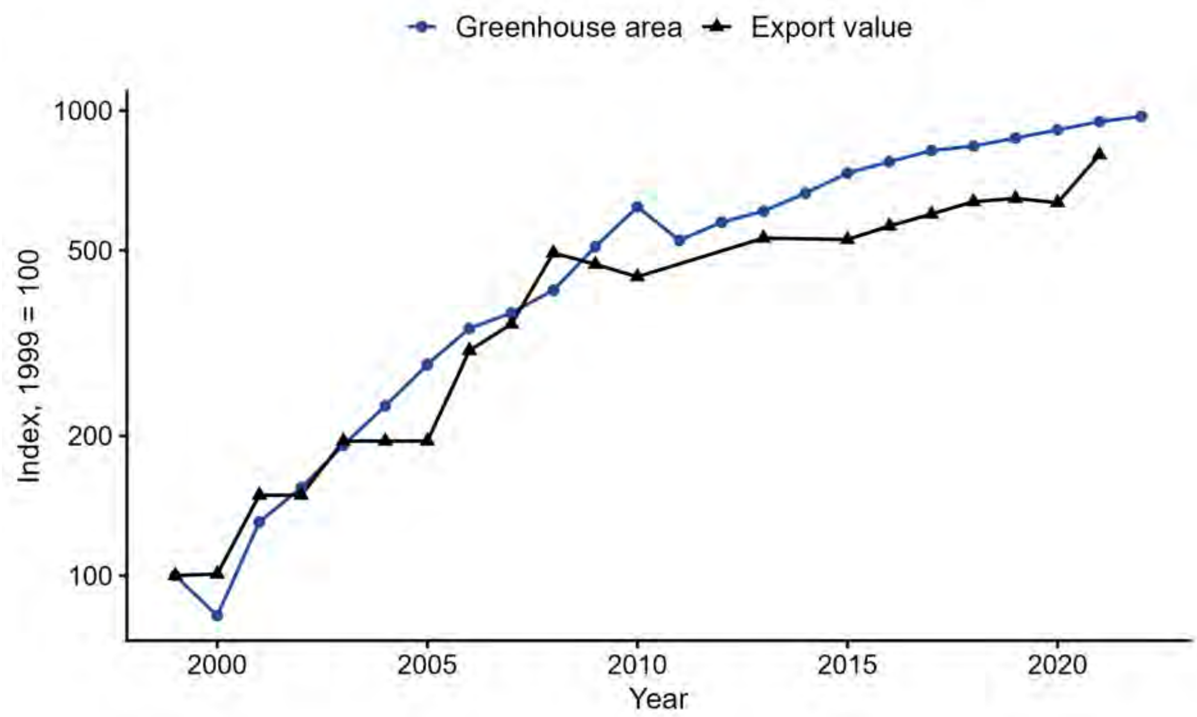
(a) 2008 **(b)** 2009 **(c)** 2010 **(d)** 2013

(e) 2014 **(f)** 2017 **(g)** 2019 **(h)** 2022

Notes: This figure displays Google Earth remotely-sensed imagery for a representative flower farm, detected in endline imagery covering the years 2021-2023. The white/silver reflective surfaces are the roofs of the greenhouses. The red contours, added manually to delineate the surface area covered by greenhouses, serve as a measure of the area under cultivation. The depicted flower farm is the same as the one shown in Figure 1. Due to cloud cover, as well as technical issues with sensors and satellites, not every location has a visible high-resolution picture available for every year. In this example, we plot all high-resolution images available between 2008 and 2022 to transparently showcase this limitation of the underlying data-generating process for measuring time of entry and area under cultivation over time.

Since we lack flower farm-level production data for the universe of flower farms identified in our sample, we are unable to test directly if our measure of greenhouse-

Figure A3: Greenhouse Area and Export Value in Kenya



Notes: This figure displays our annual measure of greenhouse area in hectares covered by a greenhouse derived from the universe of visible greenhouses in Kenya and the export value in USD PPP 2010 of cut flowers exported from Kenya for years when reported to UNCOMTRADE. To aid the graphical comparison, both time series are rescaled to a value of 100 in 1999. Export data is missing between 2010-2014. Note that the value of greenhouse area in 2021 is artificially inflated by assigning any growth in area between 2021-2023 to 2021.

covered area does in fact predict treatment intensity. We provide two pieces of suggestive evidence to support this assumption: first, Kenya Flower Council kindly shared rudimentary production data for a subset of 76 flower farms for one year (2022) with us. From this data, we can confirm that flower farm-reported hectares under cultivation correlate linearly with the treatment of interest, the number of workers employed. The ratio is 17 workers per hectare. Second, since the overwhelming majority of production is exported, we can correlate our measure of area covered by greenhouses with cut-flower exports from Kenya over time.

Figure A3 displays the resulting comparison. Our preferred measure of treatment intensity, area covered by greenhouses, tracks export value of cut flowers from Kenya to Europe remarkably closely. When rescaling both time series to a value of 100 in 1999, the relative growth in area covered by greenhouses across all of Kenya derived from our manual delineation of greenhouses over time is closely related to the annual series of export value of cut flowers leaving Kenya. The tight correlation between both variables only breaks down during the Covid-19 pandemic.

A.III Summary statistics

Table A1: Summary Statistics: All Data Sources

	N	Mean	SD	Min	Max
Panel A: Greenhouses (constructed history and 5km sublocation treatment)					
Greenhouse year of entry	237	2008.6	6.9	2000.0	2022.0
Greenhouse area, 2022 (ha)	237	40.8	210.9	0.0	3244.0
Name identified (0/1)	237	0.426	0.496	0.000	1.000
Sublocation ever treated by 2019 (0/1)	5,376	0.026	0.159	0.000	1.000
early cohort, 2000–2009 (0/1)	5,376	0.016	0.124	0.000	1.000
late cohort, 2010–2019 (0/1)	5,376	0.010	0.101	0.000	1.000
Year of first greenhouse within 5km (treated)	139	2007.5	4.8	2000.0	2019.0
Panel B: Census outcomes, 1999/2009/2019 pooled (balanced panel; women/men 15–50)					
Wage employment	4,220,450	0.204	0.403	0.000	1.000
Works for self/family	4,220,450	0.449	0.497	0.000	1.000
In-migrant (outside birth county)	4,220,450	0.286	0.452	0.000	1.000
Urban	4,219,723	0.277	0.448	0.000	1.000
Secondary enrolment (15–18)	798,494	0.344	0.475	0.000	1.000
Children ever born (women 15–29)	1,277,505	1.135	1.507	0.000	11.000
Currently married (women 15–29)	1,277,505	0.446	0.497	0.000	1.000
Births last 3 years (women 15–29)	1,271,650	0.394	0.489	0.000	1.000
Piped water (household)	1,867,161	0.269	0.444	0.000	1.000
Sewer/septic (household)	1,863,784	0.138	0.345	0.000	1.000
Electric lighting (household)	1,867,161	0.307	0.461	0.000	1.000
Clean cooking fuel (household)	1,867,161	0.125	0.331	0.000	1.000
Finished materials, 2+ of 3 (household)	1,844,525	0.498	0.500	0.000	1.000
Panel C: Night lights (sublocation)					
VIIRS-VNL night lights (max)	66,468	2.26	6.24	0.01	179.87
VIIRS-NPP night lights (max)	138,475	1.14	4.73	0.00	153.61
Panel D: FinAccess 2024 (respondent level)					
Wage employed	20,690	0.104	0.306	0.000	1.000
Flower-farm employed	20,683	0.001	0.039	0.000	1.000
Bank account	20,714	0.238	0.426	0.000	1.000
M-PESA	20,714	0.777	0.416	0.000	1.000
Mobile bank account	20,714	0.247	0.431	0.000	1.000
Chama member	20,714	0.261	0.439	0.000	1.000
SACCO member	20,714	0.112	0.316	0.000	1.000
Currently saving	20,714	0.643	0.479	0.000	1.000
NSSF pension	20,689	0.071	0.256	0.000	1.000
Ate fewer meals (1m)	20,708	0.420	0.494	0.000	1.000
Any unpaid expense (1m)	17,551	0.452	0.498	0.000	1.000
Went without medicine (12m)	20,714	0.485	0.500	0.000	1.000
School fees unpaid (12m)	20,714	0.485	0.500	0.000	1.000
Emergency: would borrow	20,609	0.144	0.351	0.000	1.000
Emergency: use savings	20,609	0.046	0.210	0.000	1.000
Emergency: work more	20,609	0.040	0.196	0.000	1.000
Emergency: cut expenses	20,609	0.005	0.069	0.000	1.000
Emergency: could not raise	20,609	0.508	0.500	0.000	1.000
Financial inclusion index	20,714	-0.000	0.587	-0.790	1.713
Savings & Pensions index	20,714	0.000	0.772	-1.343	2.186
Financial resilience index	20,714	0.007	0.693	-1.084	0.950
Emergency coping index	20,609	-0.000	0.499	-2.843	1.069
Age	20,710	39.241	17.184	16.000	105.000
Married	20,705	0.548	0.498	0.000	1.000

Notes: Summary statistics for the variables used across the paper's four data sources. Columns report the number of observations, mean, standard deviation, minimum and maximum. Panel A: the constructed greenhouse history (one observation per greenhouse) and the 5km sublocation treatment cross-section (one observation per sublocation in the balanced estimation panel); greenhouse areas in hectares. Panel B: census outcomes pooled across the 1999, 2009 and 2019 waves, on the balanced sublocation panel (present in all three waves for both sexes), women and men aged 15–50, flagged and already-treated ($\leq$1999)/post-2019 sublocations dropped; labour, migration and urban rows are person-level, housing rows household-level (deduplicated), family-formation rows women aged 15–29, secondary enrolment ages 15–18. Panel C: sublocation-level VIIRS night-light maxima on the balanced panel. Panel D: FinAccess 2024 respondents; the four indices are sign-aligned z-score composites of the listed component variables (standardized over the pooled FinAccess sample). Population and in-migrant counts, the housing/wealth PCA indices, and household primary-enrolment rates are sublocation-level constructs reported in the main tables and are omitted here.

A.IV Suitability score construction

A.IV.1 Geographic determinants of floriculture suitability

To identify the geographic determinants of flower-farm location, we conducted semi-structured interviews with managers of 18 Kenyan flower farms at the International Floriculture Trade Fair (IFTF) in Amsterdam in November 2023. We asked each manager why the farm was located where it is. The interviews revealed a clear hierarchy of siting considerations that guides the construction of our suitability score.

Natural growing conditions dominated location decisions. Altitude was the single most frequently cited determinant, mentioned in more than half of the interviews and repeatedly described as "crucial." Managers consistently viewed elevation as determining which rose varieties could be grown, with the largest-headed roses confined to the highest sites and lower-elevation farms specialising in spray or intermediate varieties. This reflects the fact that Kenyan flower farms rely on unheated plastic greenhouses that protect crops from rain, wind and pests but leave them exposed to the ambient climate.[54] As a result, the high-altitude climate remains central to commercial flower production. Cooler growing conditions lengthen the production cycle, producing longer, thicker stems and larger flower heads, the quality attributes that command a premium in export markets (Shin et al., 2023). The importance of elevation is evident even within East Africa: Ugandan rose farms, concentrated at lower elevations between Entebbe and Kampala, cannot produce large-budded roses because temperatures are too high, leaving growers confined to the less profitable small-flowered segment (Rikken, 2011).

Reliable water access emerged as another major consideration. Managers described water less as a fixed endowment than as something to be secured through boreholes, reservoirs and rainwater harvesting, but consistently emphasised that dependable water supplies were indispensable for commercial production. This accords with agronomic evidence that reliable water access and favourable climatic conditions are essential for maintaining yields and flower quality (Kenya Flower Council, 2024).

Several managers also stressed the importance of rapid road access to Nairobi's international airport, through which virtually all Kenyan cut flowers are exported.

Terrain was generally viewed as affecting production costs rather than determining feasibility. In particular, managers noted that slope matters because it facilitates irrigation. Land availability, in contrast, was uniformly described as abundant. Several managers also mentioned soil quality as a determinant of production costs rather than of feasibility: flowers can be grown hydroponically, and one manager described the ability to grow in soil as lowering costs "but not a dealbreaker". We deliberately exclude

[54] This contrasts with the heated glasshouses of the Netherlands, whose growers are largely able to insulate production from climatic shocks (Rikken, 2011).

soil nutrients from the main suitability score for a second reason: the available soil data — phosphorus, potassium and sodium from the 250 m soil nutrient maps of Hengl et al. (2017) — reflect conditions well after most flower farms in our sample entered. Because commercial floriculture itself alters soil composition through fertilizer application, irrigation and continuous cultivation, conditioning on post-entry soil measurements risks introducing a bad control that absorbs part of the treatment effect (Angrist and Pischke, 2009). We nevertheless show that the results are robust to adding the three soil-nutrient components to the score in Section A.IX.1.

Table A2: Data Sources for Suitability Score Components

Variable	Source	Native resolution	Construction
Rainfall	WorldClim v2.1 monthly precipitation climatology, 1970–2000 (Fick and Hijmans, 2017)	2.5 arc-min	Annual aggregate of monthly rasters
Solar photovoltaic output	Global Solar Atlas v2.0, long-term annual PV output (World Bank ESMAP and Solargis, 2019)	30 arc-sec	kWh per kWp installed
Groundwater depth	British Geological Survey depth-to-groundwater map of Africa, V2 (MacDonald et al., 2012)	$\sim$5 km	Modal depth class per cell, converted to class-midpoint depth (m)
Distance to water body	FAO Africover natural water bodies, Kenya, 2000 (Food and Agriculture Organization, 2000)	Vector	Distance from cell centroid to nearest water-body polygon
Elevation	WorldClim v2.1 elevation (SRTM-derived) (Fick and Hijmans, 2017)	2.5 arc-min	Indicator: cell mean elevation within p1–99 of pre-2000 greenhouse-site elevations
Slope, terrain roughness	SRTM digital elevation model	3 arc-sec ($\sim$90 m)	Computed from the DEM; slope in degrees
Distance to good road (1990s)	Africa road network, 1990s, digitized from 1:4,000,000 Michelin maps (Princeton University Library, 2021)	Vector	Good road = two or more lanes; distance from cell centroid
Electrification share (1999)	Kenya Population and Housing Census 1999, 10% microdata (KNBS)	Sublocation	Share of households with electricity as main lighting source

Notes: All gridded sources are extracted as means over 1 km^2 grid cells across Kenya, then aggregated to sublocations as area-weighted averages. Distances are computed from grid-cell centroids in UTM 37S.

Taken together, these interviews suggest that flower-farm location is shaped by three broad dimensions: climate and water, terrain, and infrastructure.

The first group, Climate and Water, includes elevation, rainfall, solar (photovoltaic) potential, proximity to major water bodies, and groundwater depth. Rather than controlling for average temperature directly, we use elevation to capture the broader high-altitude climate repeatedly emphasised by flower-farm managers. Together, these variables characterize the climatic and hydrological conditions that determine the suitability of a location for commercial flower production.

The second group, Terrain, includes slope and terrain ruggedness. These variables capture the physical suitability of the land for greenhouse construction and irrigation and therefore the cost of establishing commercial flower production.

The third group captures infrastructure. Because cut flowers are highly perishable and exported almost exclusively by air, proximity to transport infrastructure is a key determinant of production locations: farms require rapid road access to Nairobi's international airport (Button, 2020; Ksoll et al., 2023). We therefore include baseline distance to the paved road network, measured as the distance to the nearest good road (two or more lanes) in the 1990s Africa road network (Princeton University Library, 2021). We additionally include baseline electrification, measured as the share of households using electricity as their main source of lighting in 1999. Together, road access and electrification capture complementary dimensions of local infrastructure that are likely to shape both flower-farm location and subsequent local development, with road connectivity improving market access and electrification enabling modern production technologies (Asher and Novosad, 2020; Moneke, 2023). For flower farms specifically, reliable electricity is required for irrigation pumping, packhouse operations, pre-cooling, and cold-storage facilities that preserve flower quality before transport (IFC, 2020).

Table A2 summarises the variables entering each component of the suitability score, together with their data sources, spatial resolution, and construction.

A.IV.2 Construction of the greenhouse suitability score

To construct the score, we first aggregate each grid-cell covariate to the sublocation as an area-weighted average of the underlying 1 km^2 values. We then estimate an OLS linear probability model at the sublocation level, in which the dependent variable is an indicator equal to one if the sublocation lies within 5 km of a greenhouse operating at any point over our study period (2000-2019), and the explanatory variables are the geographic characteristics described above. The model is estimated on the universe of Kenyan sublocations, excluding those already within five kilometres of a greenhouse in 1999. The fitted value is the sublocation's suitability score for commercial floriculture.

The univariate regression estimates in Table A3 show how much of the variation in greenhouse presence, defined as a sublocation lying within five kilometres of a greenhouse, each component explains on its own. Most components enter with statistically significant coefficients, but their predictive power is very unevenly distributed. Elevation alone accounts for 7.1% of the variation, followed by rainfall (1.3%); distance to the road network (0.9%) and 1999 electrification (0.5%) are individually informative but explain markedly less, and no other component reaches 1%. Two implications follow. First, selection into floriculture over the 2000–2019 period we study is dominated by agroclimatic fundamentals rather than by infrastructure or terrain. Second, since no individual covariate summarises farms' location choices, we condition on economically interpretable *groups* of covariates rather than on individual predictors.

Table A3: Predictors of Greenhouse Presence within 5km

Dependent variable: greenhouse within 5km (each covariate a separate regression)	Coefficient	Std. error	Obs.	R-squared
Terrain				
Log(slope)	-0.004	(0.003)	6,182	0.000
Log(terrain roughness)	-0.035**	(0.017)	6,182	0.001
Infrastructure				
Log(distance to good road, 1990s)	-0.010***	(0.001)	6,184	0.009
Electrification share (1999)	0.095***	(0.018)	6,124	0.005
Climate and water				
Elevation (km)	-0.085***	(0.011)	6,172	0.071
Elevation (km)²	0.055***	(0.004)		
Rainfall	0.199***	(0.026)	6,172	0.013
Rainfall²	-0.116***	(0.014)		
Solar photovoltaic output	1.122	(0.765)	6,177	0.003
Solar photovoltaic output²	-0.310	(0.229)		
Log(groundwater depth)	-0.005**	(0.002)	6,164	0.001
Log(distance to water body)	-0.328***	(0.064)	6,184	0.004

Notes: OLS (linear probability) estimates. Each covariate enters a separate regression of an indicator for the sublocation lying within 5 km of a 2000–2019 greenhouse entry, in the functional form it takes in the suitability score: elevation, rainfall and solar output enter jointly with their squares (one regression per pair; R-squared reported on the linear-term row); all others enter linearly. Sample: all Kenyan sublocations not within 5 km of a greenhouse by 1999. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4 reports four multivariate specifications. Columns (1)–(3) partition the determinants of farm location into three groups: buildable terrain (column 1: slope and terrain roughness); infrastructure (column 2: distance to the 1990s road network and 1999 electrification); and growing conditions (column 3: elevation, rainfall and solar photovoltaic output, each with its quadratic term, plus groundwater depth and distance to water bodies). Column (3) does most of the predictive work: climate and water alone account for roughly four-fifths of the full model's explanatory power ($R^2 = 0.097$ of 0.124). Infrastructure explains comparatively little ($R^2 = 0.011$), while terrain explains the least ($R^2 = 0.003$), consistent with the interview evidence that land itself is not a binding constraint on where floriculture is feasible. Column (4) is the full specification that defines the suitability score ($R^2 = 0.124$), pooling all 12 components into a single fitted probability of greenhouse presence.

Table A4: Multivariate Models of Greenhouse Presence within 5km

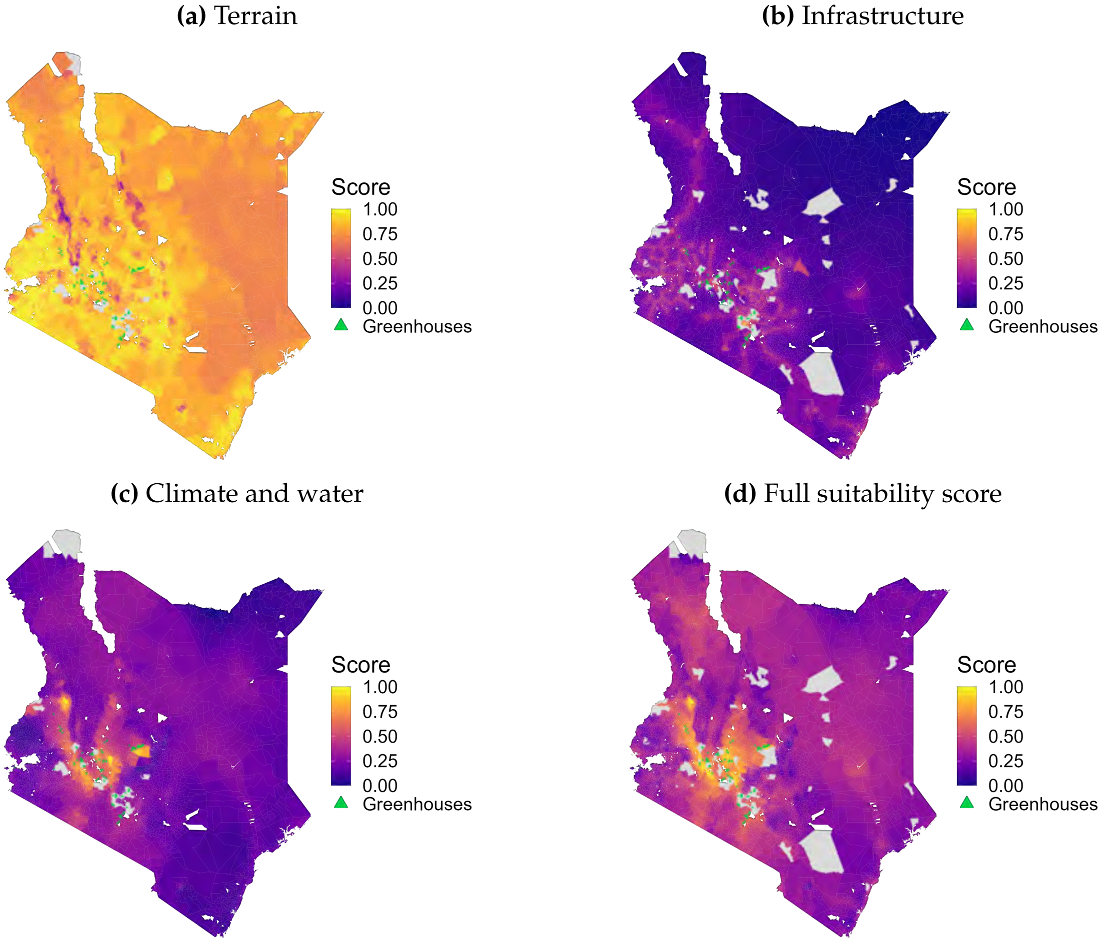
	Greenhouse within 5km			
	(1)	(2)	(3)	(4)
	Terrain	Infrastructure	Climate and water	Full
Log(slope)	0.038***			-0.004
	(0.011)			(0.014)
Log(terrain roughness)	-0.210***			-0.174***
	(0.055)			(0.061)
Log(distance to good road, 1990s)		-0.009***		-0.015***
		(0.001)		(0.002)
Electrification share (1999)		0.062***		0.007
		(0.019)		(0.019)
Elevation (km)			-0.049***	-0.040***
			(0.012)	(0.013)
Elevation (km)^2			0.049***	0.051***
			(0.004)	(0.004)
Rainfall			-0.091***	-0.165***
			(0.032)	(0.033)
Rainfall^2			0.008	0.033**
			(0.016)	(0.016)
Solar photovoltaic output			2.844***	2.242***
			(0.745)	(0.766)
Solar photovoltaic output^2			-0.829***	-0.652***
			(0.223)	(0.230)
Log(groundwater depth)			0.003	0.002
			(0.003)	(0.003)
Log(distance to water body)			-0.368***	-0.218***
			(0.069)	(0.072)
Constant	0.008	-0.078***	-4.861***	-3.444***
	(0.009)	(0.017)	(0.823)	(0.848)
Mean of dep. var.	0.027	0.027	0.027	0.027
Observations	6,182	6,124	6,151	6,091
R-squared	0.003	0.011	0.097	0.124

Notes: Linear probability model estimates; the dependent variable is an indicator for whether a sublocation first lies within 5km of a greenhouse in 2000–2019. The sample is all Kenyan sublocations, excluding those already within 5km of a greenhouse by 1999. Standard errors in parentheses. Column (4) is the full specification underlying the suitability score – the score's twelve components, the union of the three blocks. Covariates are area-weighted sublocation averages of the underlying 1km grid values. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The spatial-gradient estimation (Section A.VIII) conditions on each of the three covariate groups in turn, and then on the full score. The latter is the tightest control for selection into floriculture: it pools all 12 siting components in the proportions that best predict farm location, and enforcing overlap restricts comparisons to the agroclimatic band in which flower farms actually locate. Even so, the specifications are not strictly ordered in how much they can move an estimate: attenuation depends on how strongly a conditioning set predicts a given outcome's counterfactual trend, which varies across outcomes. For instance, if Kenya's urbanisation over this period concentrated along transport corridors, the infrastructure covariates (though comparatively weak predictors of farm location) may absorb more of the counterfactual trend in demographic outcomes than the agroclimatically weighted score. We therefore read the specifications jointly, disciplined by the placebo estimates, as a set of bounds.[55]

Figure A4 maps the four fitted scores across Kenya together with the locations of greenhouse farms entering over 2000–2019. The maps make the predictive ranking visible: the climate-and-water score sharply delineates the highland band in which farms cluster, whereas the infrastructure and terrain scores vary much more diffusely across the country — and it is the agroclimatic variation that dominates the full score.

Figure A4: Fitted greenhouse-suitability scores: covariate blocks and the full score



Notes: Each panel maps the sublocation-level fitted probability of greenhouse presence within 5 km from the corresponding specification in Table A4: column (1) in panel (a), column (2) in panel (b), column (3) in panel (c), and column (4), the full suitability score, in panel (d). Within each map, fitted values are linearly rescaled to the unit interval; the rescaling preserves the ranking and relative spacing of sublocations. Fitted values are predicted for all sublocations except those already within 5 km of a greenhouse by 1999 or with a missing score component, all of which are colored in grey. Green triangles mark greenhouse farms first operating in 2000–2019 that lie within 5 km of a mapped sublocation.

[55]The spatial gradient analysis re-estimates treatment effects at 1, 2, 5, and 10 km buffers while conditioning on this fixed 5 km score, so that only the treatment definition varies across distances. As a robustness check, we also re-fit the score separately at each buffer distance, defining greenhouse presence within 1, 2, and 10 km respectively, and conditioning on the buffer-specific score at each distance. The resulting gradient estimates are essentially identical, so we use the fixed 5 km score throughout, holding the conditioning variable constant across buffer distances.

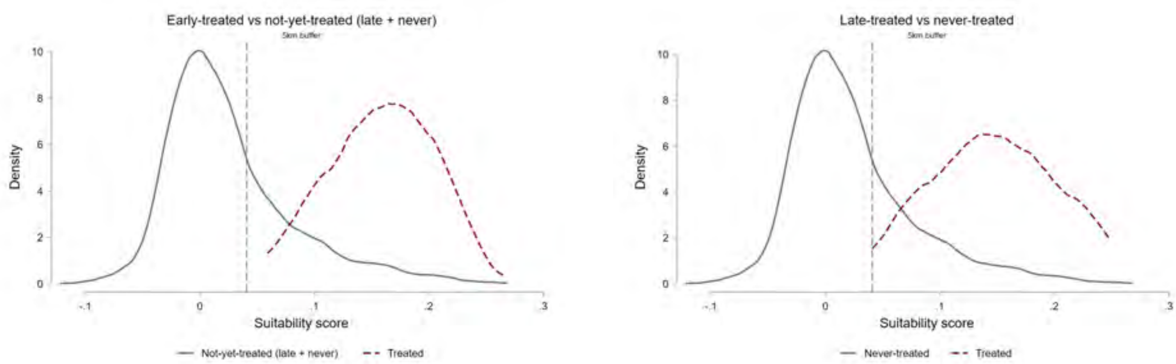
A.IV.3 Common support restrictions

Conditioning on the suitability score asks the estimator to reweight comparison sublocations to resemble treated ones. Yet large parts of Kenya are unsuitable for cut flowers: the arid north, coastal lowlands, and much of the interior score near zero, while treated sublocations are concentrated in a narrow highland band. Figure A5 shows how little comparison mass lies in the range of treated scores. Without any restriction, the estimator fits its counterfactual on comparison sublocations far outside the treated range, while the few resembling treated places receive disproportionate weight.

For effects on the treated, the overlap requirement is one-sided: we need suitable comparisons for treated sublocations, not treated counterparts for every comparison sublocation (Callaway and Sant'Anna, 2021). We therefore impose a one-sided common-support restriction (Dehejia and Wahba, 1999, 2002), excluding never-treated sublocations with scores below the *least* suitable treated sublocation while retaining all treated sublocations. We use a single cutoff, the minimum across both entry cohorts, applied before estimation. Because much of Kenya is unsuitable, this removes just over 70% of the comparison pool at the 5 km buffer. The remaining 1,469 comparison sublocations and all 139 treated sublocations form the estimation sample for every score-conditioned specification in the paper.

Figure A5: Suitability score distributions, by treatment status

(a) Early-treated (2000–2009) **(b)** Late-treated (2010–2019)



Notes: Kernel densities of the suitability score at the 5 km buffer. Each entry cohort is compared with its own comparison group; the early-versus-never comparison, visually indistinguishable from panel (a), is omitted. The dashed vertical line sits at the same score in both panels: the least suitable treated sublocation across both cohorts, whose score is the single cutoff applied to every comparison — sublocations to its left are dropped, and every treated sublocation is retained. The cutoff binds for the early cohort, which contains that sublocation; the late cohort sits entirely above it.

The treated-minimum restriction trades a large comparison pool for a more informative one: although it reduces the comparison pool from 5,212 to 1,469 sublocations, it raises its effective sample size from 191 to 289. In Appendix A.IX, Figure A18 shows that the female wage employment and population estimates are stable across alternative trimming thresholds; the Online Supplement reports the corresponding sample sizes.

A.V Changes in road connectivity

Because floriculture depends on rapid air freight to Europe, flower farm entry might plausibly pull road investment toward the export gateway, and any resulting improvement in connectivity could drive the effects we document independently of the farms themselves. We therefore test whether treated sublocations became better connected to Nairobi's international airport (JKIA). We measure travel time using a friction-surface least-cost model on a 1 km grid, combining road speeds with off-road walking at 6 km/h (Müller-Crepon, 2023). We use the 1990s Michelin/Princeton road network for the 1999 baseline (Princeton University Library, 2021) and GRIP4 (Meijer et al., 2018) and OpenStreetMap for the endline.

Table A5: Travel time to Nairobi's international airport on flower farm entry.

	Travel time to JKIA (hours)			
	Score-conditioned, comparison pool restricted		Score-conditioned, excl. Nairobi district	
	(1) Endline network: GRIP 2018	(2) Endline network: OpenStreetMap (2026)	(3) Endline network: GRIP 2018	(4) Endline network: OpenStreetMap (2026)
Flower farm entry by endline	0.343***	0.366***	0.357***	0.388***
	[0.050]	[0.053]	[0.050]	[0.053]
Never-treated mean (1999)	4.42	4.42	4.59	4.59
Early-treated sublocations (2000–2009)	84	84	84	84
Late-treated sublocations (2010–2019)	55	55	55	55
Total sublocations	1,593	1,593	1,532	1,532

Notes: Callaway and Sant'Anna (2021) difference-in-differences estimate on a 2-period balanced panel (1999 baseline, 2019 endline): travel time is measured on the baseline network and on the two endline networks, so the staggered design of the population-density tables collapses to a single two-by-two comparison, with the 2000–2009 and 2010–2019 entry cohorts pooled and never-treated sublocations as the comparison group. Treatment definition, doubly robust estimation and score-based trimming as in Table 1. Standard errors, in brackets, are clustered at the sublocation level. Nairobi district is included in columns (1)–(2) and excluded in (3)–(4), where the treated-minimum restriction is recomputed. The outcome is travel time to Jomo Kenyatta International Airport (JKIA) in hours. Endline travel time is measured on the GRIP 2018 network in columns (1) and (3) and on a 2026 OpenStreetMap snapshot in columns (2) and (4); the 1990s baseline network is identical throughout. All columns require a travel time on all three networks, so the two columns of each specification share a sample and differ only in the endline source; relative to the main score-conditioned tables this drops a handful of sublocations, none of them treated.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.VI Census data cleaning and harmonization

All census outcomes are constructed from the 10% samples of the 1999, 2009 and 2019 Kenyan censuses and harmonised across waves. We summarise the main construction choices here, with full details in the Online Supplement. Wage employment indicates working for pay for someone else. Although the activity-status question is worded identically in all three waves, in 2009 it is often inconsistent with the accompanying question on the respondent's main employer: 8.4% of women reporting an own-account employer also report working for pay, a combination that does not occur in 2019. We harmonise 2009 to the 2019 convention by reclassifying these workers as non-wage and adding them to self/family work. The employer question was not asked in 1999, so that wave is left as recorded.

Sector of occupation is available only in the 2019 census and is harmonised into seven ISIC-based groups. In 2019, an implausibly large number of workers fall under the residual "S – Other service activities" category. We reassign these workers using their free-text job descriptions, first with keyword rules and then with a curated, LLM-assisted lookup of the remaining descriptions, retaining only confident matches. In-migrants are individuals living outside their county of birth (district in 1999) who arrived within the previous decade. Housing amenities are harmonised across waves. Housing and wealth indices are constructed as within-wave first principal components of dwelling and, from 2009, asset indicators, then averaged to the sublocation and standardised to never-treated sublocations in 1999. Finally, dwelling tenure is harmonised into four groups (owner-occupied, and rented from an individual, government or other landlord), with mislabelled 1999 codes corrected using the raw IPUMS frequencies.

A.VII Supporting evidence: census-derived results

This section uses supplementary census-based evidence on employment, migration, and housing to probe our interpretation of the results in Tables 1, 2, and 3.

Table A6: Employer Decomposition 2009-2019.

	(1) Wage job, private sector	(2) Wage job, private sector (if work)	(3) Wage job, other	(4) Wage job, other (if work)	(5) Family business/farm	(6) Family business/farm (if work)
			Panel A: Women			
First census after entry	0.015**	0.023**	0.005	0.009	-0.006	-0.019
	[0.006]	[0.009]	[0.008]	[0.014]	[0.022]	[0.016]
Never-treated mean (2009)	0.040	0.066	0.079	0.135	0.494	0.771
Late-treated sublocations	55	55	55	55	55	55
Never-treated sublocations	1,469	1,469	1,469	1,469	1,469	1,469
Total sublocations	1,528	1,528	1,528	1,528	1,528	1,528
Observations	3,048	3,042	3,048	3,042	3,048	3,042
			Panel B: Men			
First census after entry	0.012	0.018	-0.010	-0.014	0.029*	0.002
	[0.012]	[0.017]	[0.013]	[0.018]	[0.017]	[0.020]
Never-treated mean (2009)	0.093	0.130	0.169	0.242	0.427	0.597
Late-treated sublocations	55	55	55	55	55	55
Never-treated sublocations	1,469	1,469	1,469	1,469	1,469	1,469
Total sublocations	1,528	1,528	1,528	1,528	1,528	1,528
Observations	3,048	3,046	3,048	3,046	3,048	3,046

Notes: Callaway and Sant'Anna (2021) staggered difference-in-differences estimates. Treatment definition and specifications as in Table 1, except that the design is 2-period (2009 pre, 2019 post): treated sublocations are the late cohort (a greenhouse arrives within 5 km in 2010–2019) and the early cohort (2000–2009) is dropped as already treated by 2009. Outcomes: (1)-(2) wage job, Private Sector employer; (3)-(4) wage job, any employer except Private Sector or State Owned Enterprise; (5)-(6) family business/farm – own family agricultural holding OR own (non-farm) family business (employment status). Even columns are conditional on working in the last 7 days. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: In-migrants and wage employment: Arrival timing relative to greenhouse entry.

	Wage job (2019)			
	All in-migrants		Came for work	
	(1) Women	(2) Men	(3) Women	(4) Men
Arrived in/after farm-entry year	0.025***	0.047***	0.113***	0.094***
	[0.009]	[0.009]	[0.014]	[0.009]
Observations	76786	68384	27346	37901
Pre-entry-arrival wage rate	0.256	0.465	0.422	0.629
Sublocations	299	298	279	292

Notes: OLS on the 2019 census cross-section (individual level), restricted to in-migrants living in treated sublocations (a greenhouse operates within 5km by 2019). The outcome is a wage job: worked for pay in the last seven days. The regressor is an indicator for having arrived in or after the sublocation's greenhouse-entry year; the omitted category is arrivals up to the year before entry. Sublocation fixed effects are included, so the coefficient compares post- vs pre-entry arrivals within the same sublocation; standard errors are clustered at the sublocation level. Cols. (1)-(2): all in-migrants with a recorded arrival year, women and men. Cols. (3)-(4): in-migrants who report work/employment as the reason for moving, women and men. Ages 15–50.

Table A8: Rise in private renting, by household migrant status.

	Share of households renting from private landlord		
	(1) All households	(2) Households with recent in-migrant	(3) Households without recent in-migrant
First census after entry	0.034***	0.033***	0.000
	[0.010]	[0.006]	[0.007]
Second census after entry	0.075***	0.052***	0.023**
	[0.017]	[0.011]	[0.011]
Decade before entry (placebo)	-0.001	-0.004	0.002
	[0.018]	[0.009]	[0.013]
Observations	4824	4824	4824
Never-treated mean (1999)	0.115	0.058	0.058
Late-treated sublocations	55	55	55
Total sublocations	1,608	1,608	1,608

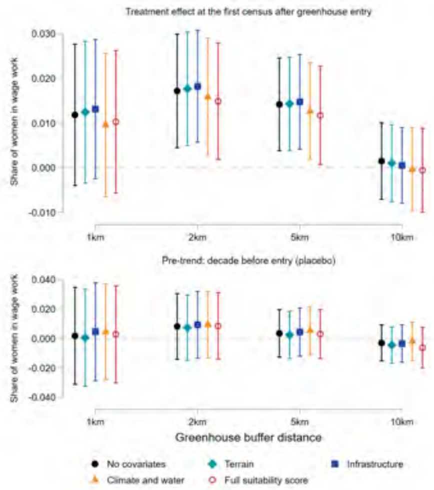
Notes: Callaway and Sant'Anna (2021) estimates; treatment definition and specifications as in Table 1. Household-level shares averaged to the sublocation; tenure harmonization and the decade in-migrant definition are detailed in the census-variable appendix. Cols. (2)–(3) split col. (1) additively by household migrant status, so (2)+(3)=(1): a household counts as migrant if any member aged 15–50 is a decade in-migrant. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.VIII Effect gradients by distance and conditioning set

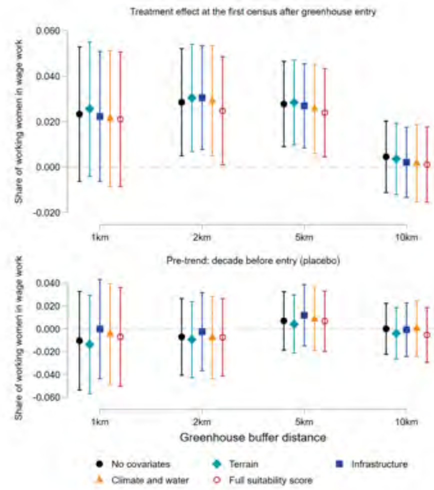
This section shows how the Callaway and Sant'Anna (2021) estimates for key outcomes in Tables 1–6 vary with distance from the farms and conditioning. At four buffer distances, we report the first post-entry treatment effect and placebo pre-trend, estimated unconditionally and conditioning on each covariate block in Table A4 or the full suitability score on the common-support sample.

Figure A6: Wage labour on flower farm entry

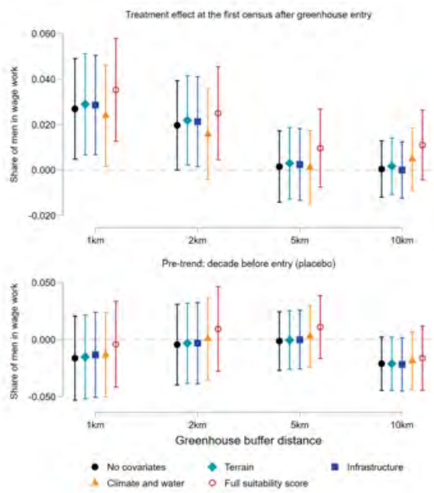
(a) Female wage labour



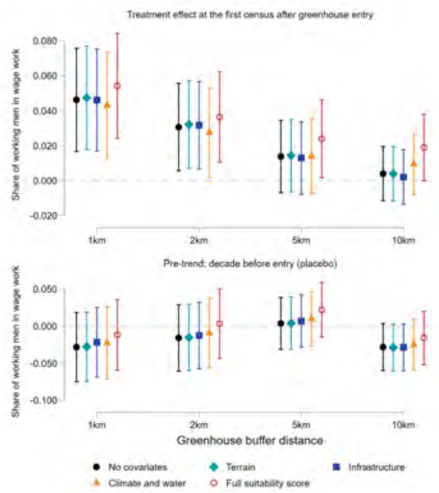
(b) Female wage labour (if working)



(c) Male wage labour



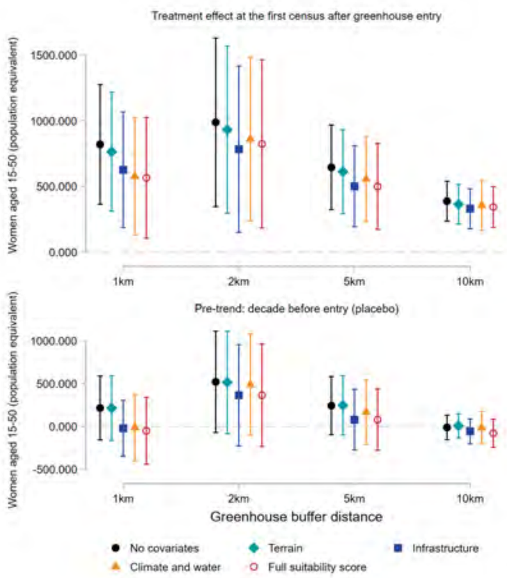
(d) Male wage labour (if working)



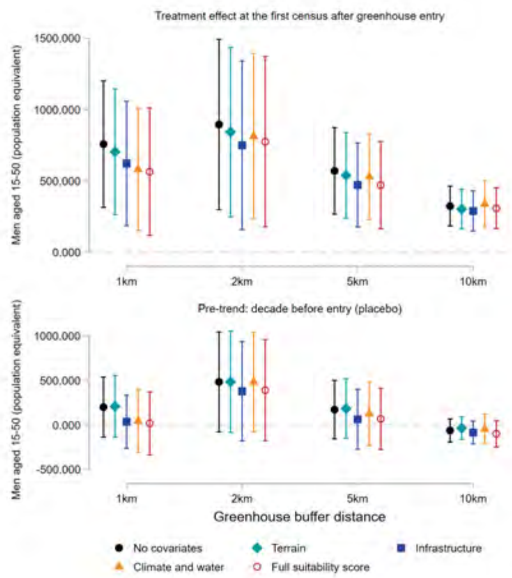
Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the outcome defined in each panel, by buffer distance. The comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score. Comparison sublocations with scores below the treated minimum are dropped. All series, including the unconditional one, are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A7: Population growth and in-migration on flower farm entry

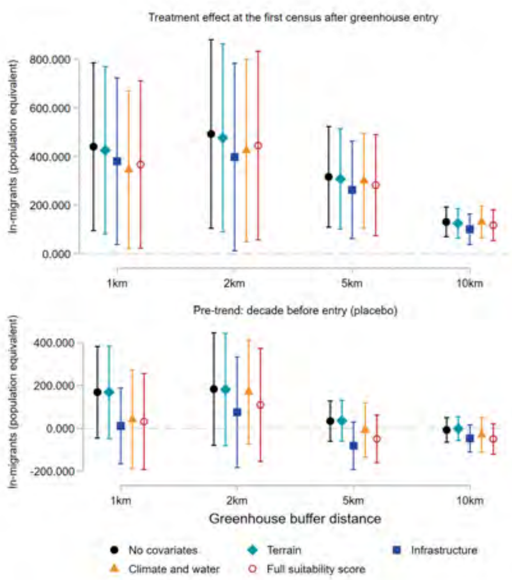
(a) Population, women aged 15-50



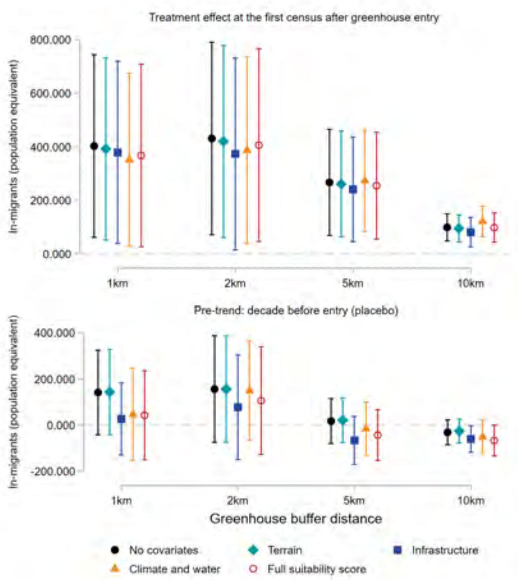
(b) Population, men aged 15-50



(c) In-migrants, women aged 15-50



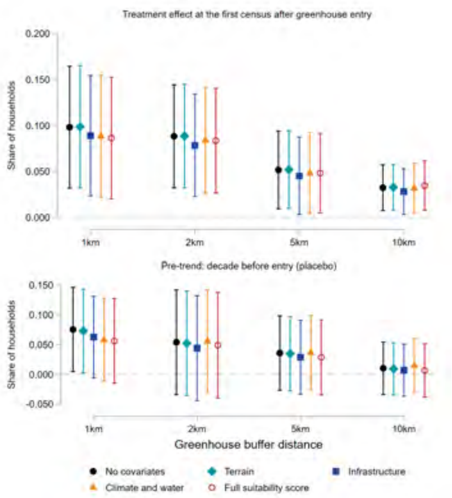
(d) In-migrants, men aged 15-50



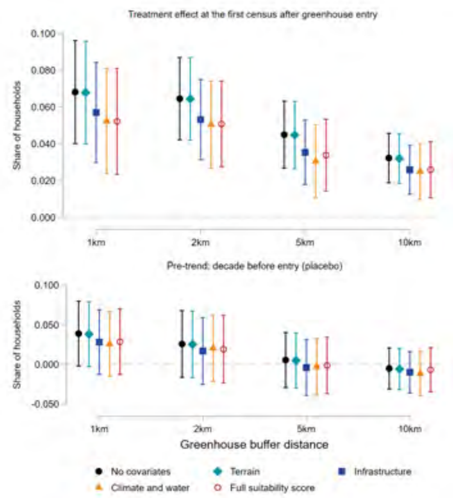
Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the outcome defined in each panel, by buffer distance. The comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score. Comparison sublocations with scores below the treated minimum are dropped. All series, including the unconditional one, are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A8: Urbanisation and private renting on flower farm entry

(a) Share of urban households



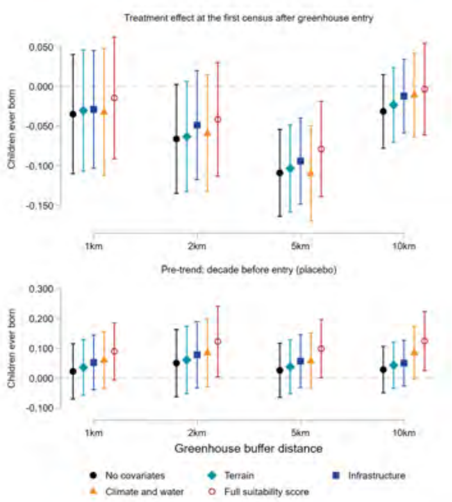
(b) Share renting from individual landlord



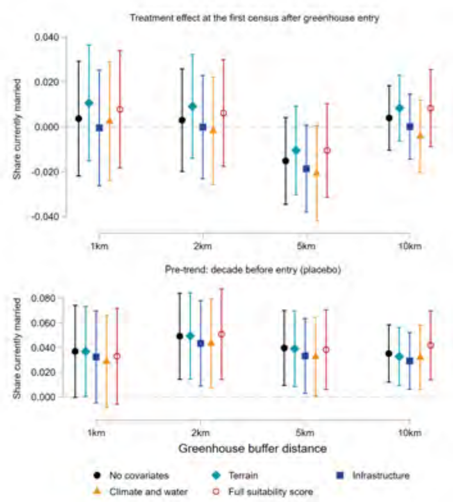
Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the outcome defined in each panel, by buffer distance. The comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score. Comparison sublocations with scores below the treated minimum are dropped. All series, including the unconditional one, are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A9: Fertility and marriage on flower farm entry, women aged 15-29

(a) Number of children ever born



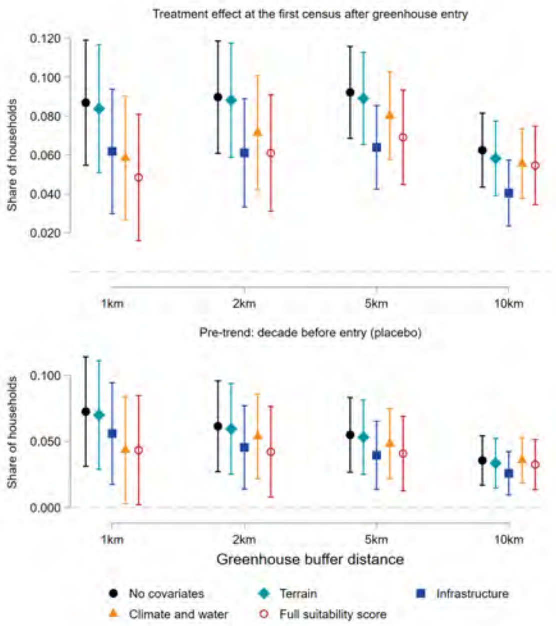
(b) Currently married



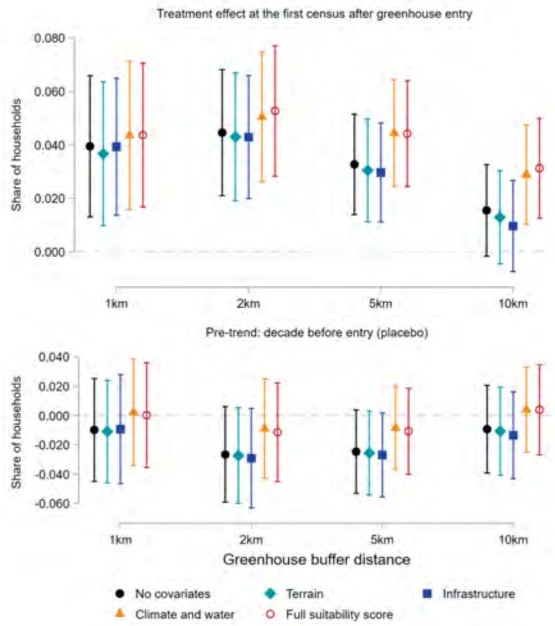
Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the outcome defined in each panel, by buffer distance. The comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score. Comparison sublocations with scores below the treated minimum are dropped. All series, including the unconditional one, are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A10: Housing quality and household wealth on flower farm entry

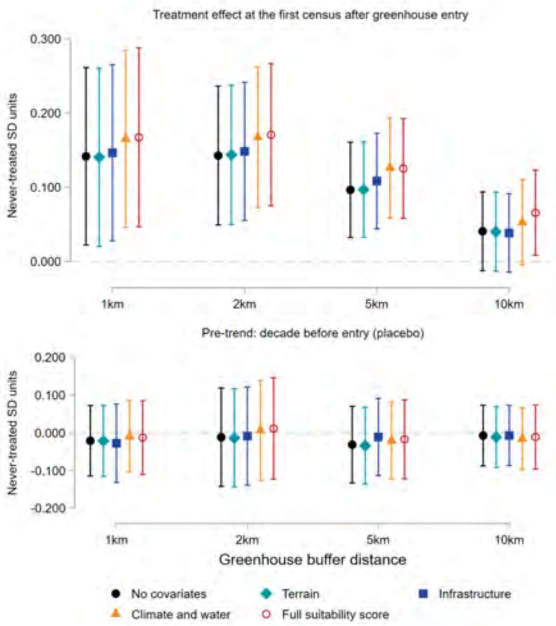
(a) Electric lighting



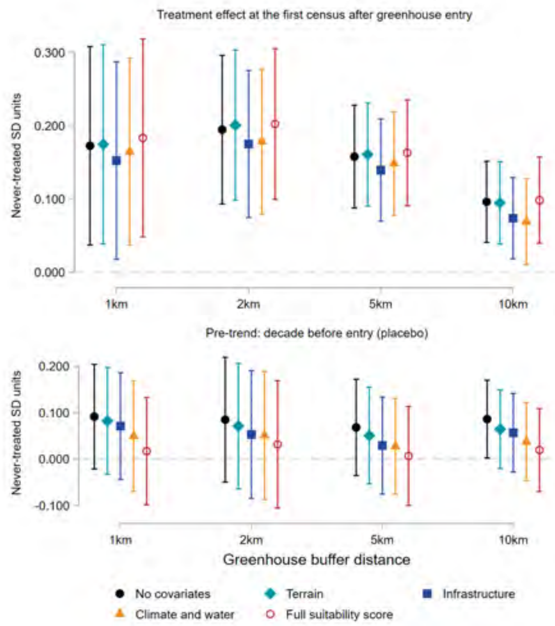
(b) Finished materials



(c) Housing index (PCA)



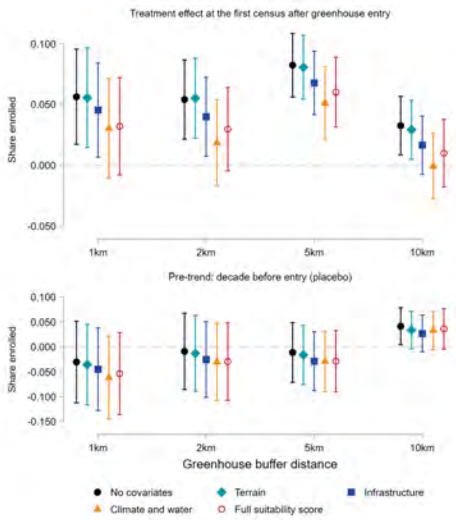
(d) Wealth index (PCA)



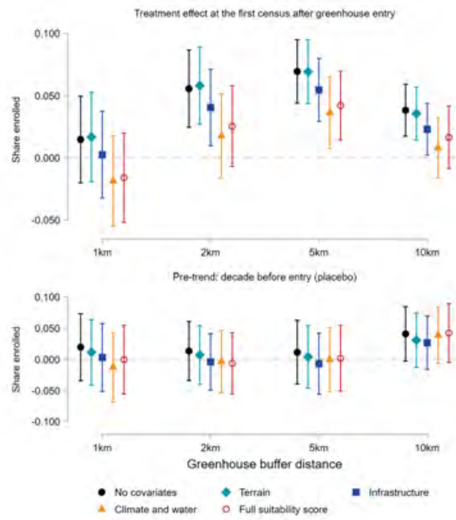
Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the outcome defined in each panel, by buffer distance. The comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score. Comparison sublocations with scores below the treated minimum are dropped. All series, including the unconditional one, are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A11: Educational enrolment and attainment on flower farm entry

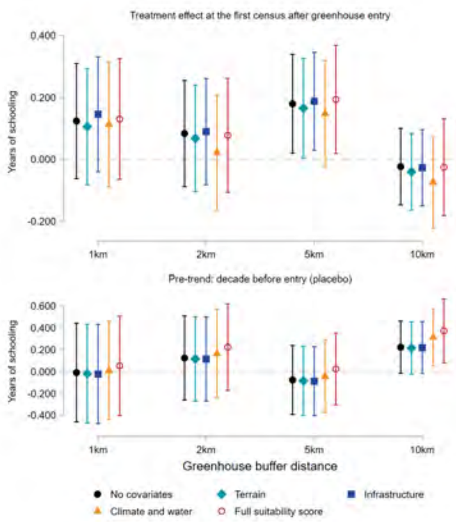
(a) Secondary school enrolment, girls



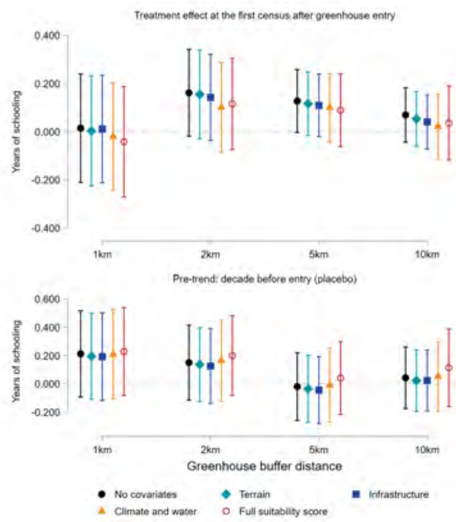
(b) Secondary school enrolment, boys



(c) Years of schooling, girls



(d) Years of schooling, boys



Notes: Callaway and Sant'Anna (2021) estimates of the effect of greenhouse entry on the outcome defined in each panel, by buffer distance. The comparison group pools never-treated with not-yet-treated sublocations. The top panel reports the treatment effect at the first census after a greenhouse enters within the buffer; the bottom panel reports the placebo pre-trend (the differential change over the decade before entry for the late-treated cohort). Five specifications are shown: (i) no covariates; and (ii)–(v) doubly robust specifications (Sant'Anna and Zhao, 2020) conditioning on (ii) buildable terrain (slope and terrain roughness); (iii) infrastructure (distance to the good-road network (1990s) and the 1999 electrification share); (iv) climatic conditions and water access (elevation, rainfall and solar potential, each with its square; groundwater depth and distance to the nearest water body); and (v) the full greenhouse-suitability score. Comparison sublocations with scores below the treated minimum are dropped. All series, including the unconditional one, are estimated on the sample of sublocations with every score component non-missing. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

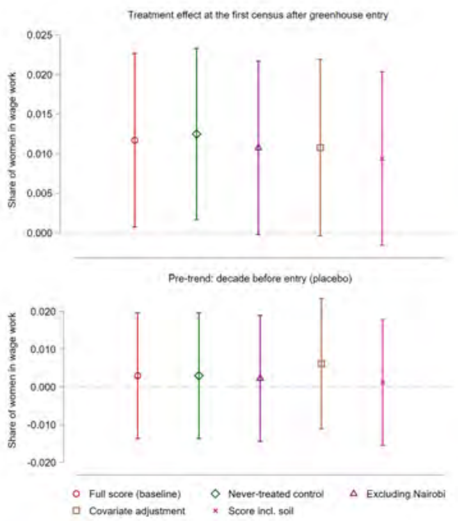
A.IX Robustness: census-derived estimates

This section reports robustness checks on the main census-based Callaway and Sant'Anna (2021) estimates. Section A.IX.1 varies the comparison group, exclusion of Nairobi, and the construction and use of the suitability score. Section A.IX.2 varies common-support restrictions from none to trimming below the treated 10th percentile for two main outcomes. Section A.IX.3 reports Honest-DiD (Rambachan and Roth, 2023) robust confidence sets for outcomes with significant placebo pre-trends, while Section A.IX.4 applies the event-study estimator of de Chaisemartin and D'Haultfœuille (2024) to alternative measures of greenhouse exposure. We show these checks for the main outcomes here; the Online Supplement reports full tables for all outcomes.

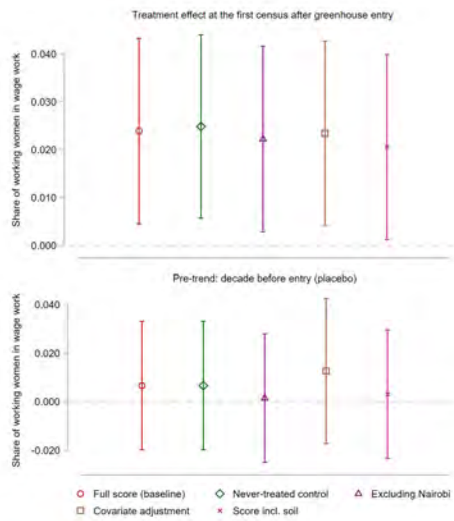
A.IX.1 Alternative specifications and samples

Figure A12: Wage labour on flower farm entry

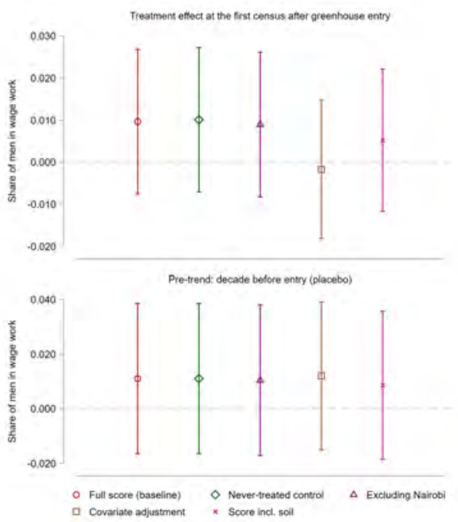
(a) Female wage labour



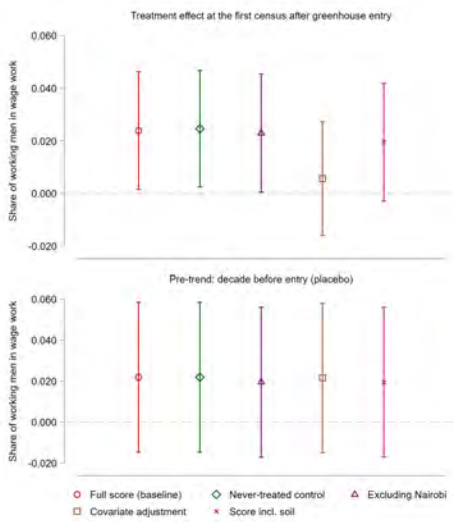
(b) Female wage labour (if working)



(c) Male wage labour



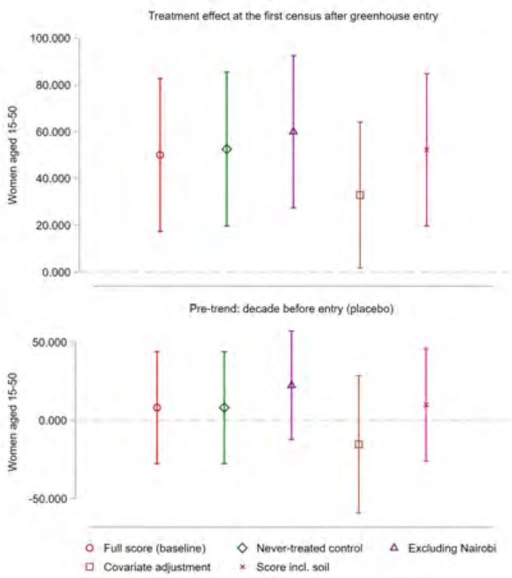
(d) Male wage labour (if working)



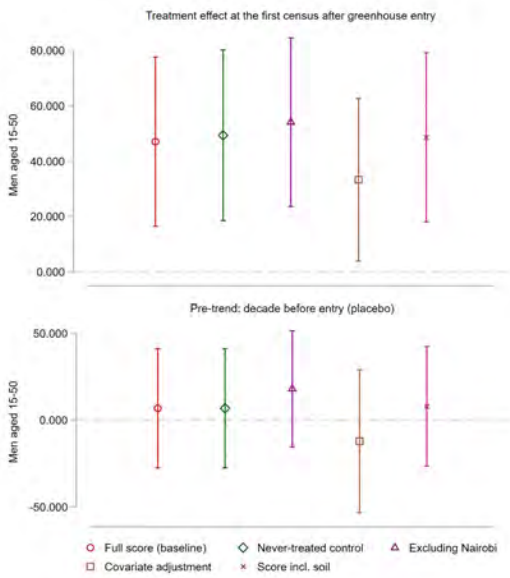
Notes: Callaway and Sant'Anna (2021) estimates at the 5 km buffer under five specifications. The estimate in red is the baseline reported in the main tables: doubly robust conditioning on the floriculture suitability score, on the common-support sample. The remaining estimates vary one ingredient at a time: a never-treated-only comparison group; excluding Nairobi district; conditioning directly on the twelve score components; and conditioning on the score refit including the three soil nutrients. The top panel reports the treatment effect at the first census after greenhouse entry; the bottom panel the placebo pre-trend over the decade before entry. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A13: Population growth and in-migration on flower farm entry

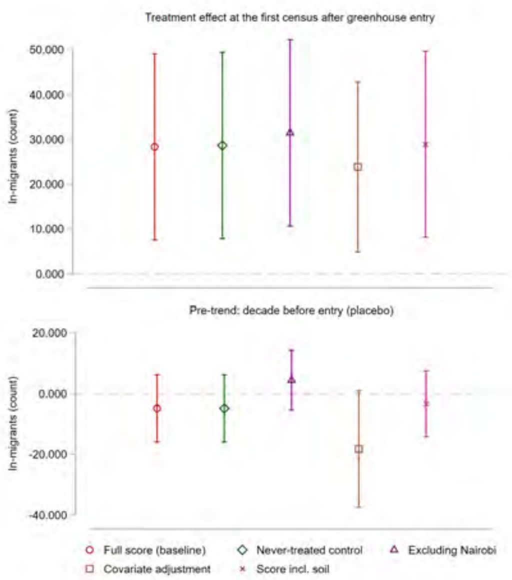
(a) Population, women aged 15-50



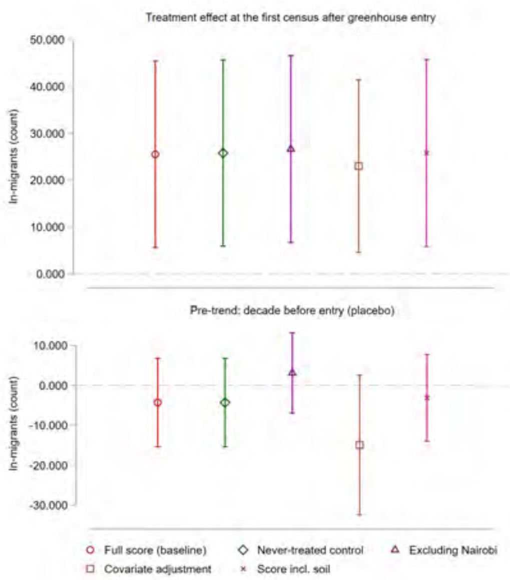
(b) Population, men aged 15-50



(c) In-migrants, women aged 15-50



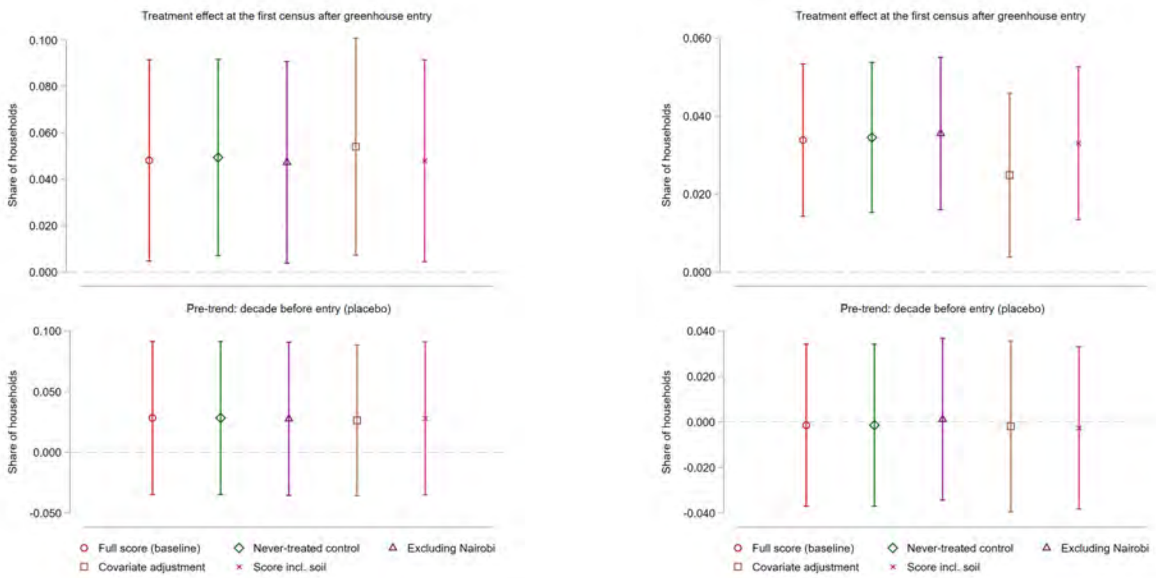
(d) In-migrants, men aged 15-50



Notes: Callaway and Sant'Anna (2021) estimates at the 5 km buffer under five specifications. The estimate in red is the baseline reported in the main tables: doubly robust conditioning on the floriculture suitability score, on the common-support sample. The remaining estimates vary one ingredient at a time: a never-treated-only comparison group; excluding Nairobi district; conditioning directly on the twelve score components; and conditioning on the score refit including the three soil nutrients. The top panel reports the treatment effect at the first census after greenhouse entry; the bottom panel the placebo pre-trend over the decade before entry. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A14: Urbanisation and housing tenure on flower farm entry

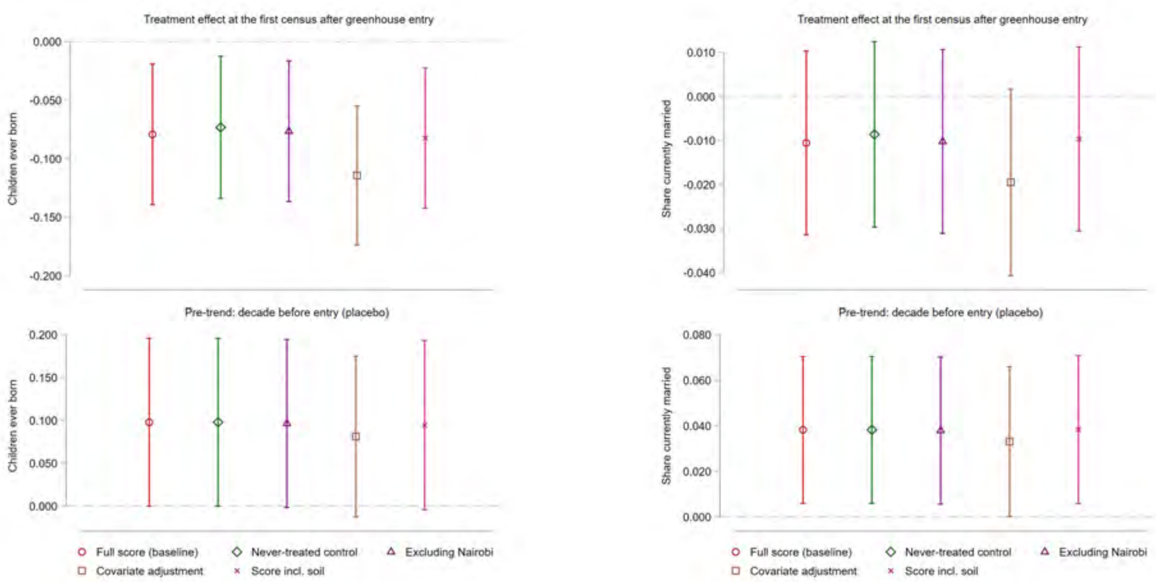
(a) Share of urban households **(b)** Share renting from an individual landlord



Notes: Callaway and Sant'Anna (2021) estimates at the 5 km buffer under five specifications. The estimate in red is the baseline reported in the main tables: doubly robust conditioning on the floriculture suitability score, on the common-support sample. The remaining estimates vary one ingredient at a time: a never-treated-only comparison group; excluding Nairobi district; conditioning directly on the twelve score components; and conditioning on the score refit including the three soil nutrients. The top panel reports the treatment effect at the first census after greenhouse entry; the bottom panel the placebo pre-trend over the decade before entry. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A15: Fertility and marriage on flower farm entry, women aged 15-29

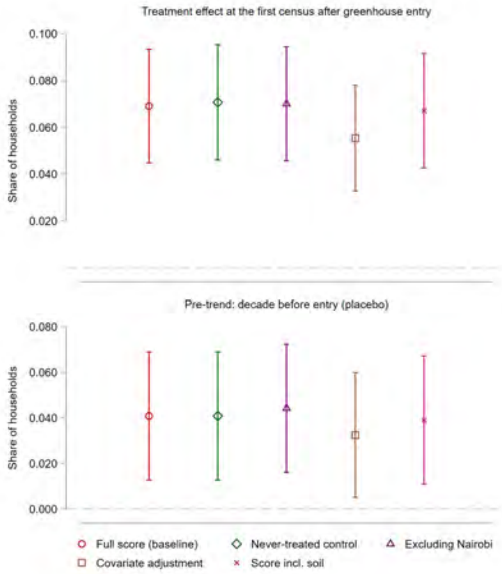
(a) Number of children ever born **(b)** Currently married



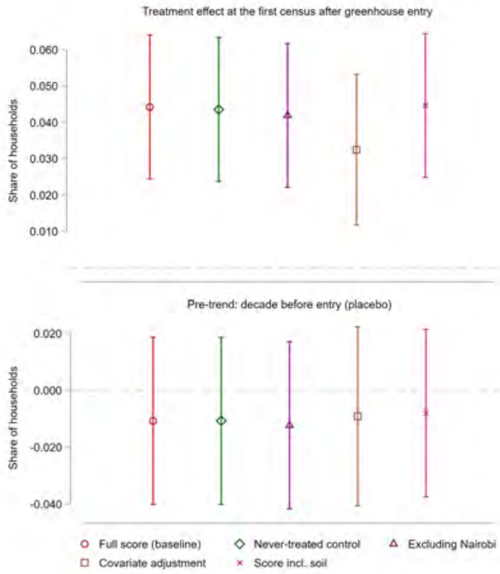
Notes: Callaway and Sant'Anna (2021) estimates at the 5 km buffer under five specifications. The estimate in red is the baseline reported in the main tables: doubly robust conditioning on the floriculture suitability score, on the common-support sample. The remaining estimates vary one ingredient at a time: a never-treated-only comparison group; excluding Nairobi district; conditioning directly on the twelve score components; and conditioning on the score refit including the three soil nutrients. The top panel reports the treatment effect at the first census after greenhouse entry; the bottom panel the placebo pre-trend over the decade before entry. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A16: Housing quality and household wealth on flower farm entry

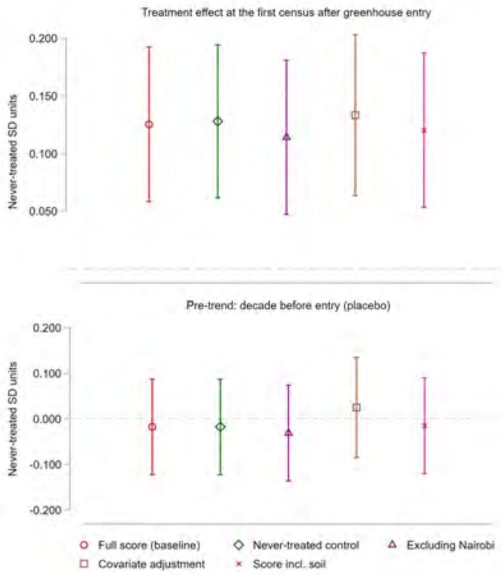
(a) Electric lighting



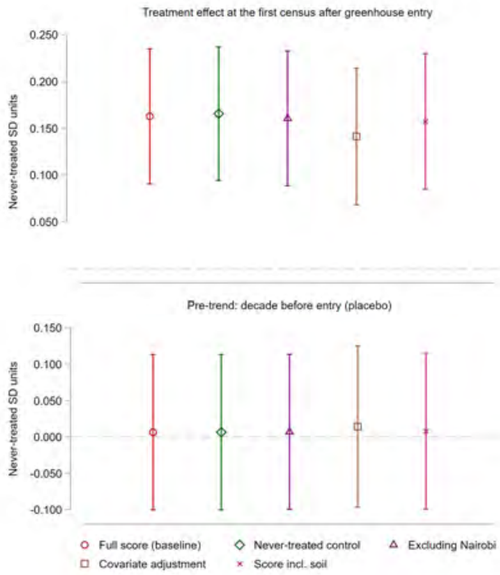
(b) Finished materials



(c) Housing index (PCA)



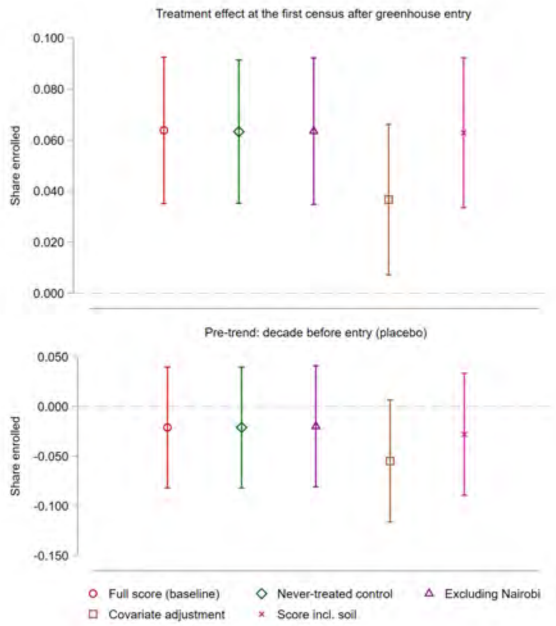
(d) Wealth index (PCA)



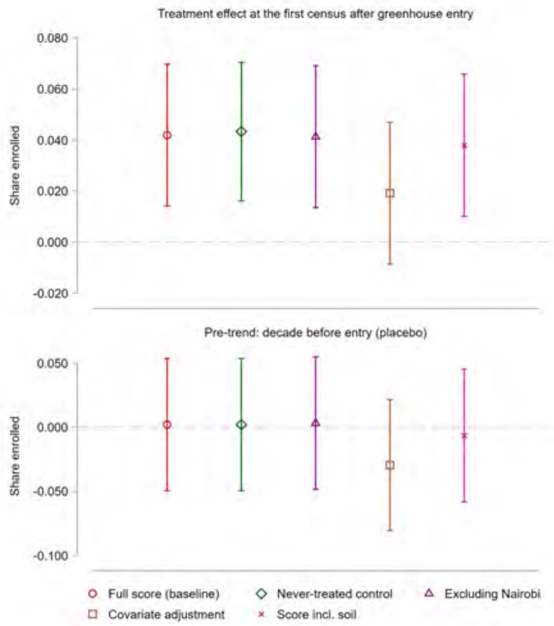
Notes: Callaway and Sant'Anna (2021) estimates at the 5 km buffer under five specifications. The estimate in red is the baseline reported in the main tables: doubly robust conditioning on the floriculture suitability score, on the common-support sample. The remaining estimates vary one ingredient at a time: a never-treated-only comparison group; excluding Nairobi district; conditioning directly on the twelve score components; and conditioning on the score refit including the three soil nutrients. The top panel reports the treatment effect at the first census after greenhouse entry; the bottom panel the placebo pre-trend over the decade before entry. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

Figure A17: Educational enrolment and attainment on flower farm entry

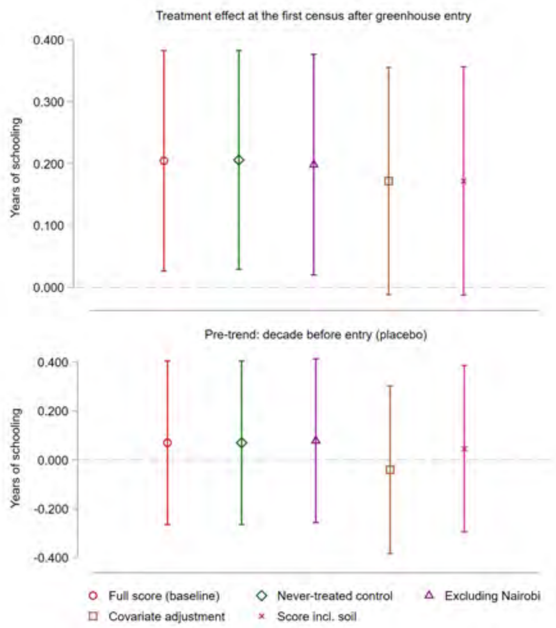
(a) Secondary school enrolment, girls (aged 14–18)



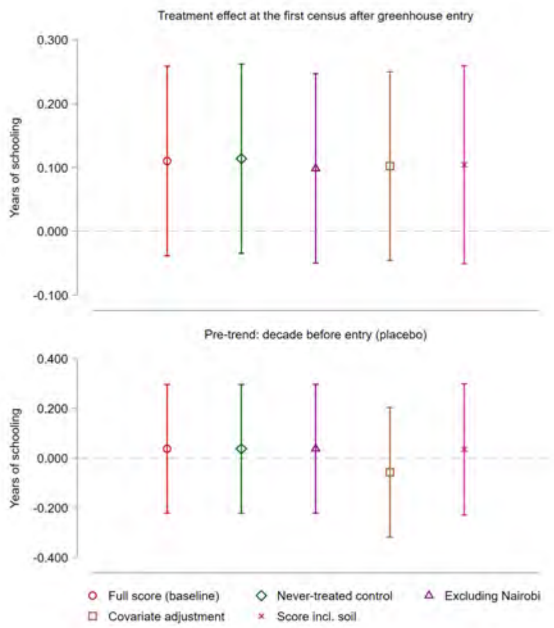
(b) Secondary school enrolment, boys (aged 14–18)



(c) Years of schooling, girls (aged 14–18)



(d) Years of schooling, boys (aged 14–18)

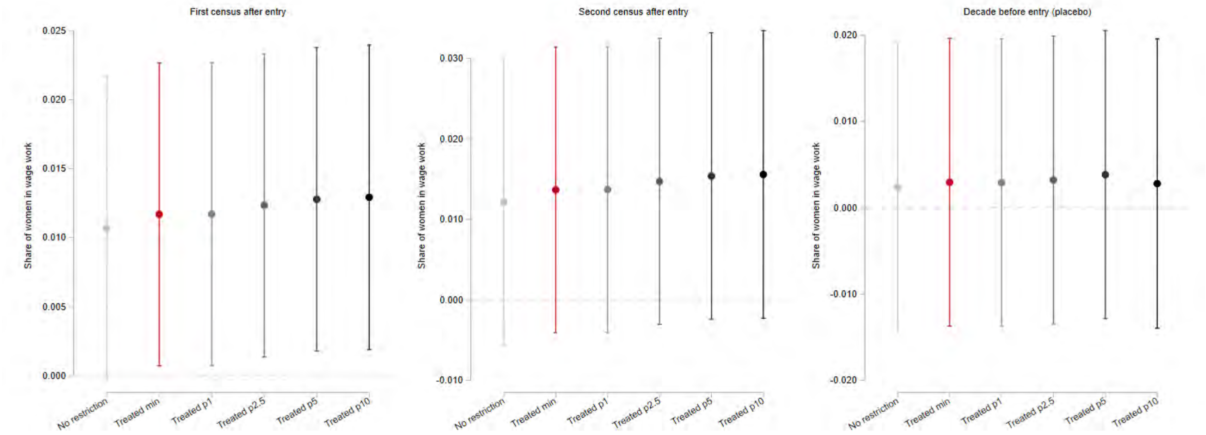


Notes: Callaway and Sant'Anna (2021) estimates at the 5 km buffer under five specifications. The estimate in red is the baseline reported in the main tables: doubly robust conditioning on the floriculture suitability score, on the common-support sample. The remaining estimates vary one ingredient at a time: a never-treated-only comparison group; excluding Nairobi district; conditioning directly on the twelve score components; and conditioning on the score refit including the three soil nutrients. The top panel reports the treatment effect at the first census after greenhouse entry; the bottom panel the placebo pre-trend over the decade before entry. Bars are 95% confidence intervals; standard errors are clustered at the sublocation level.

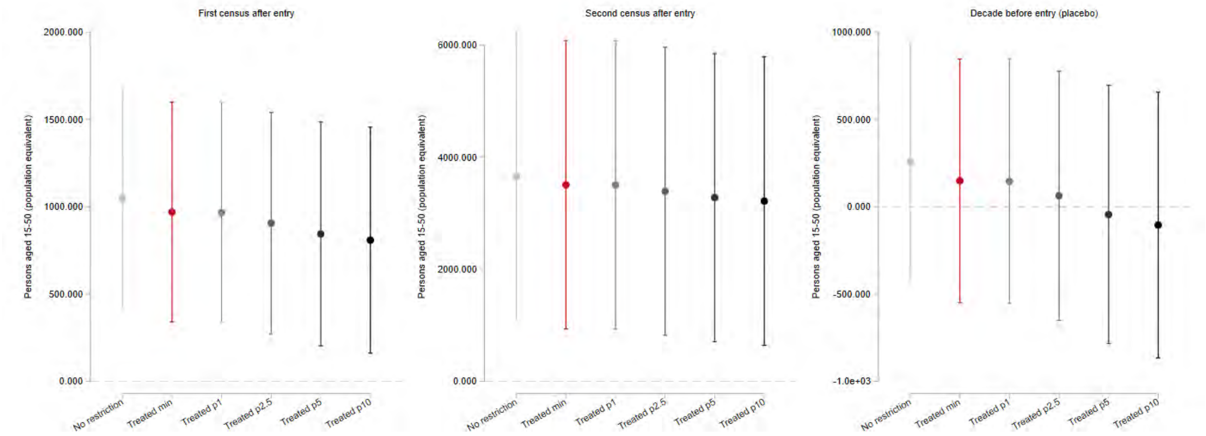
A.IX.2 Sensitivity to the trimming threshold

Figure A18: Sensitivity of the estimates to different trimming thresholds

(a) Share of women in wage work



(b) Population aged 15-50



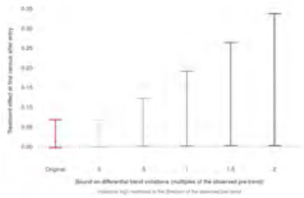
Notes: Treatment and placebo estimates at the 5 km buffer under each candidate restriction of the comparison pool, from none to dropping comparison sublocations whose suitability score falls below the treated 10th percentile (percentiles of the treated score distribution, pooled across cohorts); all treated sublocations are retained under every rule. Estimates are doubly robust (Callaway and Sant'Anna, 2021) ATTs conditioning on the suitability score, with not-yet-treated controls. Within each panel, the left plot reports the effect at the first census after greenhouse entry, the middle plot the effect at the second census after entry (2000–2009 entry cohort only), and the right plot the pre-trend placebo over the decade before entry. Markers darken as the restriction tightens; the restriction shown in red — the treated minimum — is the rule used in the main results. Bars are 95% confidence intervals.

A.IX.3 Honest-DiD sensitivity analysis

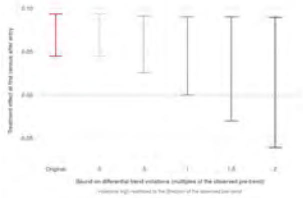
For outcomes whose placebo pre-trend is statistically significant at the 10% level, Figure A19 reports Honest-DiD robust confidence sets (Rambachan and Roth, 2023) for the first-census-after-entry effect.

Figure A19: Honest-DiD sensitivity of the first-census-after-entry effects

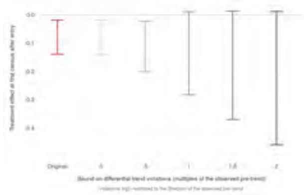
(a) Piped water



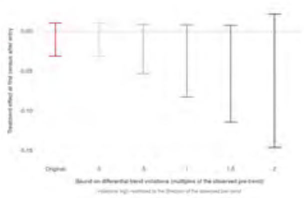
(b) Electric lighting



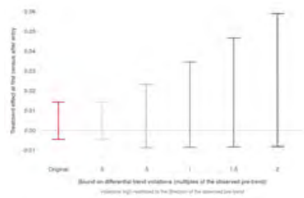
(c) Children ever born, women 15-29



(d) Currently married, women 15-29



(e) Ever given birth, women 30-44



Notes: Honest-DiD sensitivity analysis (Rambachan and Roth, 2023) for the first-census-after-entry effect, for every outcome in Tables 1–6 whose placebo pre-trend is significant at the 10% level under the baseline specification. The interval shown in red is the original 95% confidence interval; the others are 95% robust confidence sets under the relative-magnitudes restriction, allowing post-treatment violations of parallel trends up to the stated multiple of the observed pre-trend and constrained to its direction.

A.IX.4 de Chaisemartin and D'Haultfœuille (2024) estimates

Table A9: Female employment on flower farm entry

	(1) Work last 7d	(2) Wage job	(3) Wage job (if work)	(4) Self/family work	(5) Self/family work (if work)
	Panel A: Greenhouse within 5 km				
First census after entry	-0.012	0.013**	0.027***	-0.024*	-0.028***
	[0.013]	[0.006]	[0.010]	[0.013]	[0.009]
Second census after entry	0.010	0.009**	0.014*	0.001	-0.014*
	[0.008]	[0.004]	[0.008]	[0.009]	[0.008]
Decade before entry (placebo)	0.004	0.003	0.007	-0.005	-0.014
	[0.023]	[0.008]	[0.013]	[0.022]	[0.014]
	Panel B: Greenhouse area within 5 km				
First census after entry	-0.003	0.003**	0.007***	-0.006*	-0.007***
	[0.003]	[0.001]	[0.003]	[0.003]	[0.002]
Second census after entry	0.002	0.001**	0.002*	0.000	-0.002*
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Decade before entry (placebo)	0.001	0.001	0.002	-0.001	-0.003
	[0.005]	[0.002]	[0.003]	[0.005]	[0.003]
Never-treated mean (1999)	0.667	0.099	0.163	0.554	0.814
Mean greenhouse area (ha)	56.2	56.2	56.2	56.2	56.2

Notes: de Chaisemartin and D'Haultfœuille (2024) event-study estimates; treatment definitions and estimation sample as in Table 1. Score-by-census-wave interactions enter the estimator as controls in place of the doubly robust conditioning. The two panels differ only in how exposure within 5 km is measured: an indicator for any greenhouse (Panel A) and total greenhouse area (Panel B). Panel B coefficients are per 10 hectares. Standard errors, in brackets, are clustered at the sublocation level. Women aged 15–50 only; the corresponding estimates for men are in the full per-definition tables. Wage job = worked for pay in the last seven days. Self/family work = work on the household's own agricultural holding or family business. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Population growth and in-migration on flower farm entry

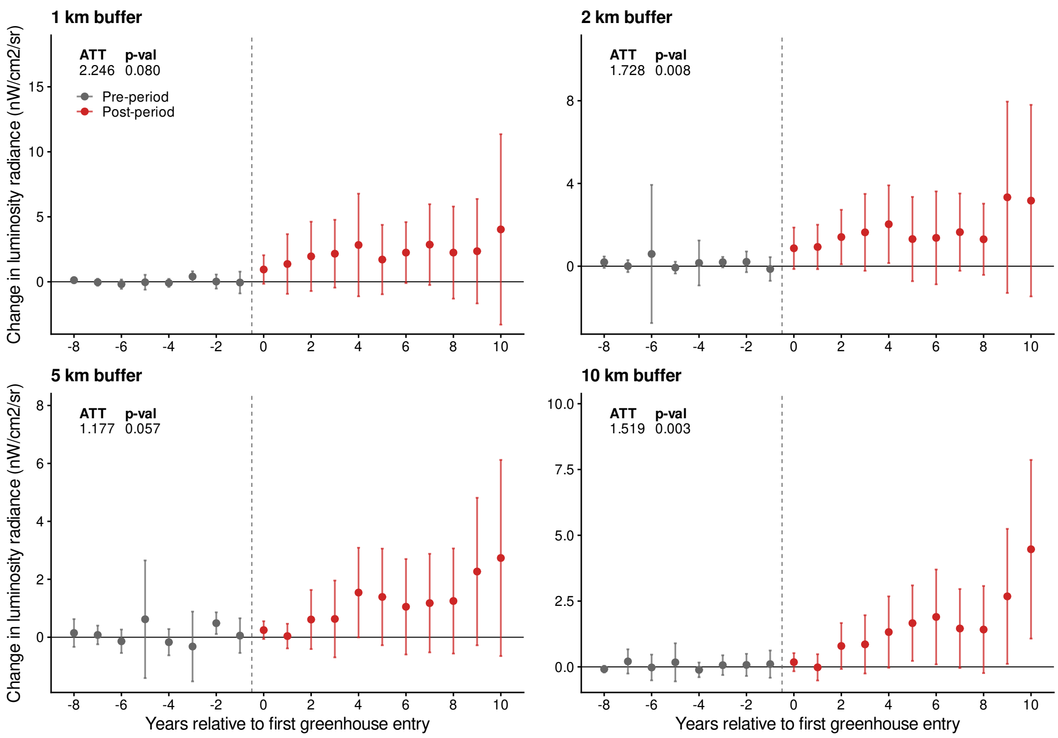
	(1) Population (total)	(2) Population (women)	(3) Female in-migrants (last 10 years)	(4) Population (men)	(5) Male in-migrants (last 10 years)
	Panel A: Greenhouse within 5 km				
First census after entry	784.6**	395.3**	236.6**	389.4**	226.1**
	[335.5]	[174.5]	[109.5]	[162.5]	[104.4]
Second census after entry	1509.1**	773.1**	488.4**	736.0**	457.0**
	[666.1]	[356.1]	[230.0]	[310.5]	[190.5]
Decade before entry (placebo)	-61.5	-34.0	-111.9*	-27.6	-92.0
	[334.7]	[171.8]	[59.8]	[164.6]	[59.4]
	Panel B: Greenhouse area within 5 km				
First census after entry	191.5**	96.4**	57.7**	95.0**	55.2**
	[81.5]	[42.5]	[26.9]	[39.4]	[25.6]
Second census after entry	245.6**	125.8**	79.5**	119.8**	74.4**
	[110.2]	[58.8]	[38.0]	[51.5]	[31.7]
Decade before entry (placebo)	-15.3	-8.4	-27.8*	-6.9	-22.8
	[85.1]	[43.7]	[15.3]	[41.7]	[15.1]
Never-treated mean (1999)	2533.1	1252.0	378.8	1281.1	406.7
Mean greenhouse area (ha)	56.2	56.2	56.2	56.2	56.2

Notes: de Chaisemartin and D'Haultfœuille (2024) event-study estimates; treatment definitions and estimation sample as in Table 2. Score-by-census-wave interactions enter the estimator as controls in place of the doubly robust conditioning. The two panels differ only in how exposure within 5 km is measured: an indicator for any greenhouse (Panel A) and total greenhouse area (Panel B). Panel B coefficients are per 10 hectares, so levels are not comparable across panels. Both panels use the same 139 switcher sublocations out of 1,612, and the same estimation sample. Standard errors, in brackets, are clustered at the sublocation level. Sublocation-level counts of persons aged 15–50, scaled by a factor of ten to population equivalents: the censuses are 10 percent samples. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.X Nighttime luminosity: alternative VIIRS measure

We re-estimate the nighttime-luminosity effects (Figure 5) using the original VIIRS annual composites (VNL) rather than the extended NPP series used in the main analysis.

Figure A20: Event studies of max. nighttime luminosity upon greenhouse arrival (2012-2023)



Notes: As in Figure 5, with one departure: luminosity is measured with the VIIRS-VNL annual composites, which are available only from 2012. The estimation window is correspondingly shorter, and sublocations treated before 2012 contribute no pre-entry observations.

A.XI FinAccess 2024 survey

The 2024 FinAccess survey is a nationally representative, geo-identified survey of financial inclusion, which we match to sublocations and our greenhouse database. We group outcomes into four indices: *financial inclusion* (bank accounts, M-PESA, mobile banking, chama and SACCO membership); *savings and pensions* (saving and NSSF contributions); *financial resilience* (avoiding missed meals, unpaid expenses, unmet medical needs and unpaid school fees); and *emergency coping* (borrowing, using savings, working more, avoiding spending cuts and raising emergency funds). Following Anderson (2008) and Kling et al. (2007), we sign components so that higher values indicate better outcomes, standardize them against pooled never-treated respondents, average the available component z-scores, and re-standardize each index to this group.

Table A11: Financial outcomes on nearby flower farm presence

Panel A: Women

	Employment		*Financial inclusion*					*Savings & Pensions*	
	(1) Wage employed	(2) Flower farm employed	(3) Bank account	(4) M-PESA	(5) Mobile bank account	(6) Chama member	(7) SACCO member	(8) Currently saving	(9) Pension (NSSF)
Greenhouse within 5km (by 2019)	0.076***	0.016***	0.092***	0.090***	0.097***	0.020	0.064***	0.120***	0.078***
	[0.017]	[0.005]	[0.018]	[0.014]	[0.019]	[0.019]	[0.017]	[0.018]	[0.016]
Observations	11783	11781	11797	11797	11797	11797	11797	11797	11785
Never-treated mean	0.074	0.000	0.188	0.770	0.213	0.320	0.088	0.628	0.045
Treated sublocs	116	116	116	116	116	116	116	116	116
Never-treated sublocs	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335

	Financial resilience				*Emergency coping*				
	(1) Ate fewer meals (1m)	(2) Any unpaid expense (1m)	(3) No medicine (12m)	(4) School fees unpaid (12m)	(5) Borrow	(6) Savings	(7) Work more	(8) Cut expenses	(9) Couldn't raise
Greenhouse within 5km (by 2019)	-0.139***	-0.077***	-0.144***	-0.101***	0.058***	0.023***	0.005	0.003	-0.111***
	[0.020]	[0.028]	[0.023]	[0.018]	[0.014]	[0.009]	[0.007]	[0.004]	[0.023]
Observations	11792	10024	11797	11797	11728	11728	11728	11728	11728
Never-treated mean	0.452	0.478	0.524	0.527	0.142	0.032	0.029	0.004	0.556
Treated sublocs	116	116	116	116	116	116	116	116	116
Never-treated sublocs	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335

Panel B: Men

	Employment		*Financial inclusion*					*Savings & Pensions*	
	(1) Wage employed	(2) Flower farm employed	(3) Bank account	(4) M-PESA	(5) Mobile bank account	(6) Chama member	(7) SACCO member	(8) Currently saving	(9) Pension (NSSF)
Greenhouse within 5km (by 2019)	0.087***	0.013**	0.114***	0.076***	0.077***	-0.014	0.038**	0.107***	0.081***
	[0.018]	[0.006]	[0.025]	[0.016]	[0.021]	[0.014]	[0.018]	[0.021]	[0.020]
Observations	8207	8202	8217	8217	8217	8217	8217	8217	8206
Never-treated mean	0.131	0.000	0.289	0.772	0.275	0.178	0.136	0.642	0.092
Treated sublocs	116	116	116	116	116	116	116	116	116
Never-treated sublocs	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324

	Financial resilience				*Emergency coping*				
	(1) Ate fewer meals (1m)	(2) Any unpaid expense (1m)	(3) No medicine (12m)	(4) School fees unpaid (12m)	(5) Borrow	(6) Savings	(7) Work more	(8) Cut expenses	(9) Couldn't raise
Greenhouse within 5km (by 2019)	-0.078***	-0.039	-0.103***	-0.109***	0.023	0.048***	0.010	-0.000	-0.112***
	[0.023]	[0.025]	[0.024]	[0.019]	[0.015]	[0.012]	[0.009]	[0.003]	[0.023]
Observations	8217	6903	8217	8217	8181	8181	8181	8181	8181
Never-treated mean	0.397	0.427	0.457	0.449	0.137	0.059	0.055	0.005	0.461
Treated sublocs	116	116	116	116	116	116	116	116	116
Never-treated sublocs	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324

Notes: Cross-sectional OLS, FinAccess 2024, respondent level; women (Panel A) and men (Panel B) estimated separately. Regressor: binary treated indicator (never-treated omitted). Controls included but not shown: urban (1999 sublocation census share), age, marital status. Standard errors clustered at the sublocation level. * p<0.10, ** p<0.05, *** p<0.01.

Table A12: Financial outcomes on nearby flower farm presence, by flower farm employment

	Financial inclusion			Savings & Pensions			Financial resilience			Emergency coping		
	All	not FF	FF	All	not FF	FF	All	not FF	FF	All	not FF	FF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Panel A: Women											
Greenhouse within 5km (by 2019)	0.311***	0.299***	1.009***	0.370***	0.344***	1.906***	0.339***	0.337***	0.437	0.193***	0.182***	0.806***
	[0.040]	[0.041]	[0.301]	[0.056]	[0.055]	[0.300]	[0.048]	[0.048]	[0.315]	[0.043]	[0.043]	[0.225]
Never-treated mean	-0.031	-0.031	-0.031	-0.058	-0.058	-0.058	-0.074	-0.074	-0.074	-0.068	-0.068	-0.068
Treated sublocs	116	116	11	116	116	11	116	116	11	116	116	11
Treated respondents	905	889	15	905	889	15	905	889	15	905	889	15
Never-treated sublocs	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335	1,335
Observations	11,797	11,781	10,907	11,797	11,781	10,907	11,797	11,781	10,907	11,728	11,712	10,838
	Panel B: Men											
Greenhouse within 5km (by 2019)	0.247***	0.240***	0.824***	0.360***	0.343***	1.801***	0.240***	0.237***	0.335	0.231***	0.230***	0.463
	[0.049]	[0.049]	[0.260]	[0.070]	[0.069]	[0.453]	[0.047]	[0.047]	[0.323]	[0.046]	[0.046]	[0.286]
Never-treated mean	0.045	0.045	0.045	0.085	0.085	0.085	0.109	0.109	0.109	0.099	0.099	0.099
Treated sublocs	116	116	7	116	116	7	116	116	7	116	116	7
Treated respondents	754	742	10	754	742	10	754	742	10	753	741	10
Never-treated sublocs	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324	1,324
Observations	8,217	8,205	7,473	8,217	8,205	7,473	8,217	8,205	7,473	8,181	8,169	7,438

Notes: Cross-sectional OLS, FinAccess 2024, respondent level. Outcomes are four z-score indices: components signed so that higher = better, z-scored against pooled (women + men) never-treated respondents, averaged over available components, and re-standardized to the pooled never-treated group (mean 0, SD 1); estimates read in never-treated standard deviations. Financial inclusion: bank account, M-PESA, mobile bank account, chama and SACCO membership. Savings & Pensions: currently saving, NSSF pension. Financial resilience: did NOT eat fewer meals (1m), no unpaid expense (1m), did not go without medicine (12m), school fees not unpaid (12m). Emergency coping: would borrow, use savings, work more, would NOT cut expenses, able to raise emergency funds. Each index reports three columns, all keeping the full never-treated group and varying only the treated group: all treated (All), non-workers (not FF), flower-farm workers (FF); treated respondents of unknown flower-farm status enter neither split, and respondents missing every component of an index are dropped. Controls included but not shown: urban (1999 sublocation census share), age, marital status. Standard errors clustered at the sublocation level. * p<0.10, ** p<0.05, *** p<0.01.

A.XII Flower farm closures

In the process of detecting flower farms from remotely-sensed imagery, we identify eleven cases of potential greenhouse closure. We manually verify each case and report the results in Table A13. We then focus on two greenhouses near Kericho town whose closure during our sample period can be independently verified.

Table A13: Cases of greenhouse closures in our data

Greenhouse ID	Location	Year of closure	Closure included	Reason
80	Mount Kenya	—	No	False positive; site disappears for some years, but reappears in later imagery.
81	Mount Kenya	—	No	
85	Tigoni	—	No	
158	Kajiado	2015	No	Limited operation (site below one or two ha) closed down within four years; possibly failed trial with negligible employment shock.
159	Kajiado	2014	No	
190	Salga/Sobea	2022	No	
10	Naivasha	2020	No	Material operation at the time of closure, but located among other large greenhouses that remained active.
160	Manyatta	2022	No	
224	Isinya	2019	No	
205	Kericho	2019	Yes	Major operation; closure confirmed by press reports. Botanical farm took over parts
206	Kericho	2019	Yes	Major operation; closure confirmed by press reports.

Note: All data from greenhouse census. All potential closure cases manually verified using Google Earth, LANDSAT and press releases.

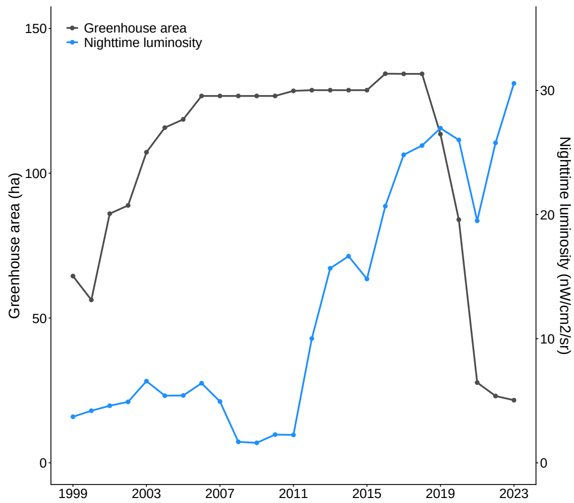
Figure A21: Kericho greenhouse closures 2004-2024

(a) March 2004

(b) March/April 2014

(c) March 2024

(d) Greenhouse area and nightlights



Notes: The first three panels depict Google Earth Engine remotely-sensed imagery of Kericho town and its surroundings. Three flower farms are visible in the very South-West, very South-East and East, whereas the town is visible in the North (centre of image). High quality images of the area exist from 2004 onward. Panel (a) depicts a small and panel (b) a much denser town while the South-Eastern and Eastern flower farms expand. Panel (c) shows two of the three greenhouses having closed, the Southwestern one fully reforested, whereas the Southeastern one is in partial use by another firm growing botanicals. Panel (d) overlays our data on greenhouse area under cultivation for all three greenhouses in the picture with maximum simulated VIIRS (1999-2023) nighttime luminosity readings, masking the spatial extent of all three flower farms from the nighttime lights measure.

For additional results and details about our data, please find the **online supplement** to this paper under the following link: Online Supplement