Trac(k)tors of Change: Monitoring, Tractor Mobility and Agricultural Mechanization in Kenya*

Sophie Nottmeyer†
CEMFI

November 2025

Click here for the latest version.

Abstract

Limited access to mechanization constrains agricultural productivity in developing economies, where rental markets for capital goods remain thin. This paper studies how supply-side monitoring frictions shape the spatial allocation of capital in tractor rental markets in Kenya. I evaluate the introduction of a GPS tracking application that enables tractor owners to monitor operators remotely, combining high-frequency GPS data from around 1,200 tractors and nearly one million georeferenced fields with satellite imagery, an original farmer survey, and a quantitative spatial model. Following adoption, monitored tractors gradually expand their range of operations and reallocate toward areas with higher returns to mechanization, consistent with reduced monitoring costs and improved capital allocation. Remote sensing evidence shows that fields visited by monitored tractors exhibit more effective immediate land preparation and greater sustained vegetation growth than comparable nearby fields, suggesting productivity gains at destination. Model estimates indicate that digital monitoring reduces spatial misallocation by 15% and raises aggregate output by 2%, demonstrating the potential of digital technologies to enhance market efficiency in settings involving the delegated operation of mobile capital.

*I thank Mónica Martínez-Bravo, Diego Puga, and Paula Bustos for their guidance and support. I also thank Giorgio Pietrabissa, Manuel Arellano, Dávid Krisztián Nagy, Elisa Giannone, Pierre-Philippe Combes, Nick Tsivanidis, David Atkin, Kevin Donovan, Julieta Caunedo, Namrata Kala, Mushfiq Mobarak, Mark Rosenzweig, Lauren Bergquist, Nicholas Ryan, Kyle Emerick, Tillmann von Carnap, and others for their helpful comments. I am grateful to the team at Hello Tractor, especially Susan Njihia, for their trust and for generously sharing their data with me. I gratefully acknowledge financial support from STEG-CEPR, the Weiss Fund, the Spanish State Research Agency (MDM-2016-0684), and the European Union (ERC ClimateAdapt, Project 101088060). All errors are mine.

†Centro de Estudios Monetarios y Financieros (CEMFI), Casado del Alisal 5, 28014 Madrid, Spain. E-mail: sophie.nottmeyer@cemfi.edu.es. Website: sophienottmeyer.github.io. First version: October 2025.

1 Introduction

Agricultural productivity remains persistently low in many developing countries, particularly in Sub-Saharan Africa, which is commonly attributed to a lack of adoption of modern technologies such as mechanization. For most farmers, individual ownership of costly equipment such as tractors is not viable, due to both financial constraints and their inability to capture returns to scale (Foster and Rosenzweig, 2022). Rental markets could, in principle, help small producers overcome these capital indivisibilities and expand access to mechanization, as recently shown in urban manufacturing clusters in Uganda (Bassi et al., 2022). Why, then, are they not more widespread in agriculture?

This paper examines how monitoring frictions on the supply side of tractor rental markets limit the efficient spatial allocation of tractors. In these markets, tractor owners typically rely on hired operators to drive their tractor and provide mechanization services to farmers. However, owners cannot easily observe operators' actions, giving rise to moral hazard: operators have incentives to deviate from the owners' instructions, for example, by underreporting the number of jobs or acres they worked on.[1] To mitigate these risks, owners prefer to keep tractors in areas where direct supervision and control (e.g. through unannounced field visits or checks on fuel consumption upon return) are easier: near their own residence or in places where they can rely on trusted contacts for oversight. As a result, moral hazard effectively limits the geographic reach of individual tractors and the set of locations to which owners are willing to send them.[2]

Monitoring frictions can have important implications for the spatial scope and efficiency of tractor rental markets. Depending on the initial distribution of tractors, these individual constraints to tractor mobility can aggregate into significant spatial misallocation, leaving remote but productive regions underserved. Two features make monitoring frictions particularly salient in this context. First, rapid urbanization has given rise to urban 'absentee' owners (e.g. civil servants or entrepreneurs who invest in tractors as income-generating assets), so tractors are not necessarily based where they are the most productive (Kaumbutho and Takeshima, 2023). Second, demand for tractor services is highly seasonal and staggered across regions, so even if tractors were already based where they are needed, efficient use would require them to move as the rains shift.[3]

[1] Other dimensions along which incentives may be misaligned include overstating the distance travelled, misusing fuel, driving carelessly to save time, or taking shortcuts on poor roads, all of which increase costs for owners.

[2] Although this paper focuses on agricultural mechanization in a developing-country setting, similar monitoring problems arise in other sectors involving the delegated operation of mobile capital, such as trucking, ride-hailing, and public transportation, across both developing and advanced economies.

[3] In his seminal work on the introduction of onboard diagnostic computers in the US trucking industry,

To study the importance of these monitoring frictions, I leverage the introduction of a new digital monitoring technology for tractors in Kenya. I use the universe of GPS activity records from tractors that adopted this technology, which provide high-frequency information on exactly where tractors operate. To complement these data, I conducted an original farmer survey that informs the analysis and validates key outcome measures. Combining the GPS data with satellite imagery and other spatial data, I provide micro-level empirical evidence that digital monitoring reduces spatial frictions, improving the allocation of tractors and increasing agricultural productivity. Building on these findings, I develop and estimate a quantitative spatial model of tractor location choice, calibrated with tractor flows constructed from the GPS records, to quantify the aggregate efficiency gains from reducing monitoring frictions.

The new monitoring technology I study consists of a GPS tracking device installed in the tractor and a mobile/web application that allows owners to monitor operators remotely. Introduced by the Kenyan start-up Hello Tractor (HT) in 2018 and sold directly to tractor owners, adoption has been rapidly increasing. By 2025, around 1,225 tractors have adopted the technology, representing between 1.73% (Census 2019) and 8.75% (FAO 2008) of all tractors in the country. The app provides detailed information on the number, location and size of jobs, as well as the routes taken and distance travelled, mitigating the moral hazard problem.

The analysis mainly draws on unique tractor GPS data obtained from the app provider, containing information on both tractor activities and home locations, which allows me to study the spatial allocation of tractors relative to their base. Each observation corresponds to a single job serviced by a tractor, defined by the exact geographic boundary of the serviced area, a timestamp, and a unique tractor identifier. The data covers the universe of jobs, comprising about 900,000 observations from 2018 to 2025 across the entire country. I complement this data with a small original survey of farmers in one county to guide my analysis, and combine it with different gridded spatial data sources matched by location.

I provide two pieces of empirical evidence that the new monitoring technology reshaped the spatial allocation of tractors and improved agricultural productivity. First, I use within-tractor variation to study how job locations evolve after adoption, taking the first month—when owners and operators are still adapting to monitoring—as a baseline for comparison. I find that tractors progressively expand their range of operations, working on jobs farther from their base, on average about 55 km after one year and 112 km

Hubbard (2000) similarly distinguishes two motives for adoption: mitigating moral hazard between firms and drivers (an incentive motive) and improving resource allocation and dispatch decisions (a coordination motive).

after two years (an increase of about 80% and 160% of the baseline mean, respectively). These patterns are consistent with a reduction in monitoring-induced spatial frictions. Matching job locations to potential yield estimates from the FAO-GAEZ database, I further show that tractors shift toward areas with higher marginal returns to mechanization, indicating a more efficient spatial allocation.

Second, I show that this reallocation translates into real productivity gains at the field level. Using an event study design, I estimate the effect of a monitored tractor visit on a proxy for agricultural productivity derived from satellite data, matching each job to a nearby field with similar characteristics and pre-trends in the outcome. Fields visited by a monitored tractor exhibit a significantly larger increase in vegetation growth in the year of the visit. Since the control group may also have mechanized through other tractors, this effect can be interpreted either as a lower bound on the productivity gains from mechanization or as the differential impact of new or more effective mechanization due to monitoring.

Based on these insights, I develop a quantitative spatial model of tractor location choice to quantify the aggregate gains from monitoring, accounting for equilibrium effects of reallocating scarce capital across space. In the model, tractor owners choose where to send their tractors given their base location and distance-based moving costs that incorporate the monitoring component (owner-operator moral hazard). The new monitoring technology effectively reduces the elasticity of these costs with respect to distance: for a given distance, owners perceive the cost of moving to be lower when monitoring is possible.

I estimate the location choice parameters using bilateral tractor flows constructed from the GPS data, introducing a new, more direct approach to identify the elasticity of tractor flows (the Fréchet shape parameter) that uses behavioral responses to exogenous variation in potential yields as a shifter of returns. Solving the model given the observed initial distribution of tractor capacity across locations then allows me to recover the relative attractiveness of locations by revealed preference.

Using the estimated model, I conduct counterfactual analyses to quantify the aggregate gains from improved monitoring. The main comparative statics exercise differs from standard applications that vary trade costs or travel times (e.g. through transport infrastructure improvements). Instead, I vary the elasticity of moving costs with respect to distance, structurally estimating parameter changes based on the same empirical variation in job locations driven by gradual adaptation to monitoring documented before.

First, I simulate the allocation of jobs among HT tractors under different levels of monitoring experience, comparing an equilibrium corresponding to the early stages of adoption to one where all adopters behave as if they had five years of experience (fully adapted). I find that monitoring reduces spatial misallocation of tractors by 15%, raising aggregate output by 2%. Second, I extend this analysis by rolling out the monitoring technology to all tractors in Kenya. Finally, I consider an alternative policy the government could use to encourage owners to travel farther, a fuel subsidy calibrated to deliver the same efficiency gains, and show that it is substantially more costly than digital monitoring.

Overall, this paper demonstrates how monitoring frictions on the *supply* side of tractor rental markets can constrain the mobility of productive capital, limiting its efficient spatial allocation. Combining unique tractor GPS data with an original farmer survey, satellite data, and a quantitative spatial model, I provide new evidence that digital monitoring can alleviate these frictions, enabling tractors to reallocate more efficiently across locations and generating substantial productivity gains. The paper also highlights the potential of digital technologies as scalable and cost-effective tools to improve access to mechanization for smallholder farmers, supporting broader processes of structural transformation in Kenya and beyond.

Related Literature. This paper contributes to several lines of existing work. First, it relates to the literature on barriers to agricultural technology adoption, recently reviewed by Suri and Udry (2022) for Sub-Saharan Africa. While most of this literature emphasizes demand-side barriers such as credit and information constraints, a smaller but growing body of work highlights input market access as a key limitation. For example, Aggarwal et al. (2024) show that limited access to fetilizer markets constrains adoption in Tanzania. Agricultural mechanization, a key driver of structural transformation and productivity growth, has received renewed interest within this context. In particular, Caunedo and Kala (2021) randomly subsidize farmers' access to productive capital in India and document positive returns to mechanization, empasizing demand-side constraints. This paper instead focuses on supply-side barriers that determine farmers' access to mechanization across locations and highlights the potential of new digital technologies to reduce transaction costs.

Despite its importance, agricultural mechanization in Sub-Saharan Africa remains understudied largely due to limited data availability.[4] This paper is the first to use GPS tracking data in general, and the data obtained from Hello Tractor in particular, to pro-

[4]In the case of Kenya, the last nationwide Census of Agriculture was conducted in 1979; the next is planned for 2025-26, with a pilot currently underway. No official data source, including government agencies, systematically records tractor ownership or use, and FAOSTAT discontinued its global time series on agricultural machinery in 2009.

vide new insights on the barriers to agricultural mechanization. It also demonstrates how to measure spatially disaggregate agricultural output using satellite data, without relying on large-scale field data collection.[5]

Second, this paper relates to the literature on capital rental markets and misallocation (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2017). Bassi et al. (2022) are the first to highlight the importance of rental markets in reducing capital misallocation (in urban manufacturing clusters in Uganda). Relative to their work, this paper emphasizes the spatial dimension–specifically, frictions that prevent rental markets from working efficiently when supply and demand are geographically separated and capital is mobile. The paper also relates to Walker et al. (2024), who study the macroeconomic consequences of input indivisibility when sharing is technologically impossible. In particular, my paper similarly examines how frictions affect aggregate outcomes, when sharing is feasible but constrained. In the context of agricultural mechanization, Caunedo et al. (2022) examine government-run equipment hiring centers in India and the distributional consequences of alternative dispatch technologies. In contrast, this paper studies a decentralized, privately operated rental market in Kenya, where baseline mechanization rates are substantially lower.

A parallel literature studies the diffusion of tractors, highlighting the role of private rental markets in facilitating this process. This work spans both the historical experience of now-developed economies such as the US (Olmstead and Rhode, 2001; Manuelli and Seshadri, 2014) and more recent growth episodes in China (Yang et al., 2013) and Myanmar (Belton et al., 2021), where migratory service provision across regions played a central role. This paper provides new evidence on the early stages of this process in Sub-Saharan Africa.

Third, this paper also relates to the empirical literature on monitoring and firm performance, beginning with seminal work by Hubbard (2000, 2003) and Baker and Hubbard (2003), who study how the introduction of onboard diagnostic computers affected the US trucking industry. With the recent rise of digital technologies, a growing number of papers studies the impact of monitoring in the developing country context (de Rochambeau, 2021; Houeix, 2025). In particular, Kelley et al. (2024) conduct a randomized experiment with commuting minibuses in Nairobi, showing that owners can use monitoring tech-

[5]In related work, I collaborate with remote-sensing scientists to predict crop yields from satellite data using machine learning techniques, where the scarcity of labeled training data again poses a major challenge. To address this, we show that transfer learning, pre-training an algorithm on abundant US data to learn general associations and fine-tuning it to the Kenyan context using county-level crop yield statistics, can overcome this limitation (Bohra et al., 2025). Thus far, the yield estimates are at the county level, but we are currently working on disaggregating them to the pixel level to generate annual crop-yield maps for Kenya for public and research use.

nologies to mitigate moral hazard with drivers by inducing higher effort and reducing risk-taking, thereby increasing firm profitability. My paper complements this work by linking firm-level moral hazard to market-level efficiency across locations.

Finally, this paper contributes to a growing literature at the intersection of development and spatial economics, which applies spatial tools to study development questions (recently reviewed by Bryan et al. (2025) and Storeygard (2025)). In particular, recent work analyzes spatial mobility in low and middle income countries using large-scale digital data, such as smartphone app data and call detail records (Kreindler and Miyauchi (2023) for Sri Lanka and Bangladesh; Blanchard et al. (2025) for Kenya, Nigeria and Tanzania). My paper demonstrates how a quantitative spatial model can be adapted to a distinct context, studying the mobility of capital rather than humans, using novel GPS tracking data. It also connects to recent evidence by Kreindler et al. (2025), who document that workers in Nairobi are willing to accept lower pay to avoid unfamiliar neighborhoods, highlighting how limited prior exposure constrains spatial mobility. Analogously, my paper shows how digital monitoring technologies can reduce this bias by enabling tractors to operate in unfamiliar but more productive areas.

The remainder of the paper is structured as follows. Section 2 describes the institutional setting. Section 3 introduces the data. Section 4 presents empirical evidence on the effects of the new monitoring technology. Section 5 introduces the quantitative spatial model of tractor location choice. Section 6 discusses its estimation and calibration. Section 7 undertakes the counterfactuals. Section 8 concludes.

2 Institutional Setting

This section outlines key institutional details. First, I briefly discuss the economic importance of agricultural mechanization and its current state of adoption in Kenya. Second, I describe the organization of private tractor rental markets. Then, I discuss the moral hazard problem underlying the monitoring friction and how it interacts with distance. Finally, I describe the new digital monitoring technology.

2.1 Agriculture and Mechanization in Kenya

Agriculture plays a key role in the Kenyan economy, accounting for 21% of GDP and 34% of total employment (World Development Indicators 2019). Beyond the official employment estimates, nearly half of the the population (46%) engages in crop farming, mostly for subsistence - 76% of household produce primarily for their own consumption (Census 2019). The sector is dominated by small, fragmented farms, with a median size of just one acre, divided into one or two plots (Kenya Integrated Household Budget Survey, KIHBS, 2015/16).

As in much of Sub-Saharan Africa, agricultural productivity remains very low, largely due to limited adoption of productive technologies such as mechanization, irrigation, and other inputs like fertilizers. The number of tractors in use in the region has stagnated since the 1960s, in contrast to other developing countries in Asia for example, where tractor use has increased significantly since the 1980s (Sims et al., 2016). In East Africa, as of 2008, only 17% of the cultivated area was prepared using tractors, while 50% still relied on manual labor and 32% on draft animal power (Mrema et al., 2018).

Data on mechanization in Kenya is limited, but available evidence suggests that tractor ownership and usage are similarly low. Fewer than 1% of households own a tractor (Census 2019), while around 23% of farmers report expenditures on 'Tractor/Oxen Plough' and 5% spend on 'Hire of Machinery' (KIHBS 2015/16).

Tractor use is both appropriate and beneficial in terms of productivity, primarily through labor savings and improving the timeliness of farming operations.[6] Maize, Kenya's primary crop, is well-suited to mechanization. While ploughing a one-acre plot takes a tractor approximately 1.5h, doing so by hand can take a farmer nearly a week. Given that most farming in Kenya is rain-fed, mechanization also enhances the timeliness of planting, which is crucial for maximizing yields.

While farmers generally recognize these benefits, their ability to use tractors is constrained by availability. Investment in and ownership of tractors is limited by the small scale of production for most producers, preventing them from realizing economies of scale. As a result, farmers access mechanization almost exclusively through rental markets, which appear active here given the difference between ownership and use noted above.

2.2 Private Tractor Rental Markets

I study private tractor rental markets where individual owners provide mechanization services, including an employed operator, to farmers. A key feature of this market is that tractor owners are disproportionately located in urban areas, complicating their ability to effectively manage and monitor tractor usage in rural areas, where tractors are needed.

This study focuses on the decentralized private tractor rental market, which differs from government-run schemes or specialized hiring centers that operate from fixed loca-

[6] In principle, mechanization also enables farmers to cultivate larger areas, but frictional land markets limit expansion in this context. In addition, tractor ploughing may also offer agronomic benefits, such as improved soil quality.

tions. In this market, individual owners - typically with single tractors or small fleets of four-wheel tractors with 50-75 hp - offer mechanization services to farmers. The range of mechanization services depends on the corresponding implements they own, but ploughing is by far the most commonly demanded service. Rentals include a trained operator, who is employed by the owner and assigned to a specific tractor, as farmers usually do not operate the tractors themselves. Owners primarily rely on personal connections, referrals from previous customers, and, in some cases, local intermediaries to secure clients. Rental prices are typically charged per acre, covering both service and transport costs, with payments commonly collected by the operator on the day of service. For ploughing, rates range from 3,000-5,000 KES (approximately 25 USD) per acre, depending on the region.[7]

It is important to note that tractor owners are not necessarily farmers themselves, who rent out excess capacity. In fact, 43% of tractor owners do not engage in agricultural production (Census 2019). Instead, the majority of them are local entrepreneurs who invest in tractors to provide mechanization services as a business or to supplement income from other employment and ventures. As wealth is accumulating in cities due to rapid urbanization, those with the capital to invest in tractors are increasingly located in urban areas. In Kenya, 33% of tractor owners live in cities, with Nairobi having the highest concentration of tractors (8.9%; Census 2019). This creates a spatial mismatch between where capital is owned and where it is most productive.

2.3 Monitoring Friction

In this context, moral hazard between tractor owners and their operators, arising from owners' inability to observe operators' behavior, limits the locations to which owners are willing to send their tractors.

The key challenge in this market is that tractor owners rely on their operators to provide services to farmers, but they cannot observe output - specifically, how many and which plots or how many acres were serviced overall - making it difficult to assess the revenue they should expect to receive from their operators at the end of the day. Operators are typically paid based on output and often retain a share of revenue as their wage. However, since output cannot be verified by the owner, operators have an incentive to shirk and underreport earnings without facing consequences. Additionally, they may engage in actions, such as servicing unauthorized jobs or stealing fuel, which increase operating and maintenance cost, all of which are borne solely by the owner.[8]

[7] This cost corresponds to up to 64% of the average monthly consumption expenditure per adult equivalent (KIHBS 2015/16).

[8] Moral hazard problems of this kind are not unique to tractor rental markets, but they have also been

Focus group discussions and qualitative interviews with tractor owners reveal the extent of moral hazard, with most owners reporting difficulties in controlling their operators, leading to financial losses. As one particularly frustrated owner expressed, *"You are the servant, they are the boss. You own the tractor, but they make more money than you."* To mitigate these risks, owners employ various strategies, such as unannounced field visits, hiring third parties to check on operators, requiring tractors to return to their base every night, and tracking fuel consumption levels and mileage upon return. However, these 'in-person' monitoring efforts are costly and require tractors to remain nearby. Given that many tractor owners are often located in urban areas, this suggests that the spatial allocation of mechanization may be inefficient.

2.4 New Monitoring Technology

To study the role of monitoring frictions, I exploit the introduction of a new monitoring technology that allows tractor owners to track their tractors remotely, improving their ability to manage operators to behave in their interest.

The monitoring technology consists of two components: a GPS tracking device and a mobile/web application. The GPS device is installed on tractors by a professional mechanic and sends real-time location signals with 2-meter accuracy to an online server through the mobile network every 5 seconds ("ping").[9] The accompanying application gives tractor owners access to the data (see Figure 1), including real-time location and daily activity reports (left) summarizing the number of jobs serviced, acres ploughed, working time, and distance traveled. Additional features include a movement replay view that allows owners to visualize the tractor's exact movements over time (right), an area measurement tool and geo-fence alerts.[10]

The technology was introduced by the local start-up Hello Tractor (HT) in 2018, as a digital monitoring solution specifically tailored to tractors.[11] Owners pay a one-time installation fee of 125 USD and annual subscription of 60 USD per user to access the app.[12] It was marketed through dealer and NGO networks (e.g. Farm to Market Alliance),

documented in other industries in low-income countries, such as transportation or trucking.

[9] The device uses an international SIM card that can roam across telecom networks. If tractors move to an area with low coverage, it stores data offline for up to 30 days and automatically uploads it once reconnected. The device also has an independent internal battery, allowing it to continue to send pings when the tractor is idle or powered off, although at a lower frequency.

[10] Geo-fence alerts notify owners when a tractor leaves a designated area defined by the owner. Premium features include fuel consumption monitoring and remote immobilization, which allows the owner to immobilize the tractor remotely when it is doing unauthorized work or is in an unapproved location.

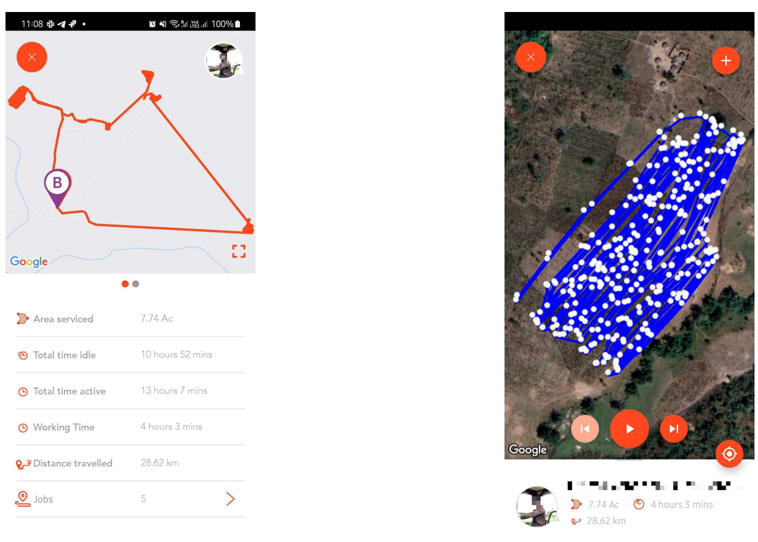
[11] https://hellotractor.com/

[12] Pricing as of 2023. For reference, the average tractor works on around 4 acres per day during the season and charges around 25 USD per acre, so the cost of adoption is roughly equivalent to two days worth of

Facebook campaigns, and demonstration events, with key messages such as *"Technology for smarter, better maintained and more profitable tractors."* and *"Track tractor and operator performance to ensure maximum machine uptime, profits, and reduce fraud."* In principle, GPS tracking devices with basic software are available in Kenya since around 2015, but they are not designed for tractors and do not offer the same visibility and flexibility as HT's solution.

Giving tractor owners the ability to track their tractors reduces the moral hazard problem by enhancing their visibility into tractor activity (e.g. detecting unauthorized jobs), thus improving their signal about revenue and allowing them to better align operators' incentives with their own interests. While the way owners use the monitoring tool to manage their operations may vary, operators generally are aware they are being monitored, but do not know where exactly the device is installed and cannot access the app to prevent tempering. Anecdotal evidence from interviews with HT customers suggests that owners find significant value in using the monitoring technology to mitigate moral hazard. In fact, experienced users report high levels of trust and delegation to their operators once the monitoring friction is resolved.

Figure 1. Tractor Owner Mobile Application



(a) Daily Activity Report **(b)** Track Replay

Note: Screenshots showing two different views within the Hello Tractor mobile application for tractor owners. (a) Summary view displaying all serviced jobs and key metrics for a given day. (b) Detailed view of a single job visualizing the exact movements of the tractor working on a field, where white dots represent individual GPS pings.

work.

3 Data

My main data source is GPS tracking data from tractors using the new monitoring technology, covering the entire period since its introduction. A key advantage of this data is that it allows me to observe tractor use at a very high spatial and temporal resolution. In this section, I explain how tractor activity records were generated and describe how I identify tractor home locations. Secondary data sources are discussed in the Appendix A.

3.1 Activity Records

The tractor activity records I have access to are an aggregated version of GPS pings generated for the daily activity reports in the app. Specifically, this consists of work area polygons, which were created by drawing the boundary around clusters of back-and-forth movements when connecting pings chronologically. I call my unit of observation a 'job', defined as a single polygon with its associated tractor ID and timestamp.

The final dataset contains information on around 917,000 jobs that were serviced by 1,225 unique tractors over the time period 2018-2025. This corresponds to 1.73% of tractors in the 2019 Census and 8.75% of tractors in Kenya last reported in FAOSTAT in 2008 (Mrema et al., 2018). The total area serviced, aggregated over the entire time period while avoiding double counting, accounts for around 14.5% of total cropland[13], a significant share given that only around 17% of cultivated area is estimated to be mechanized in East Africa (Mrema et al., 2018).

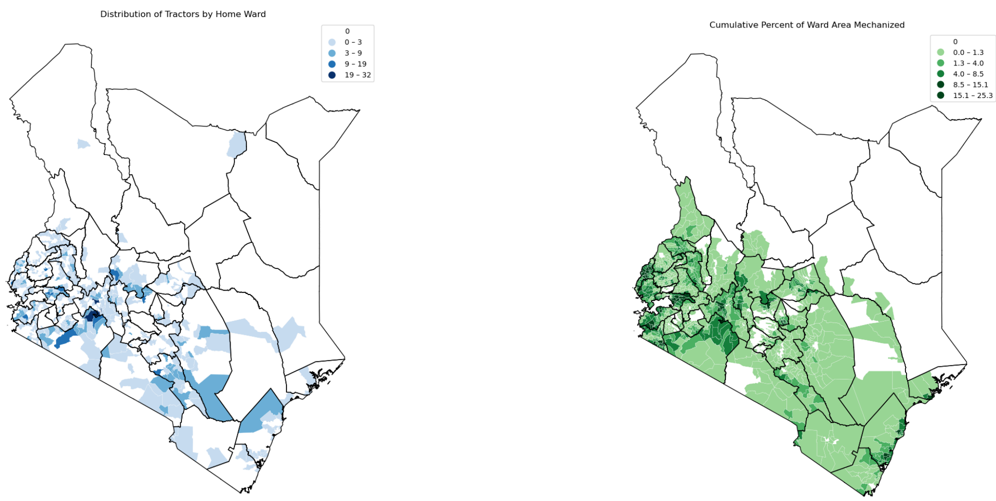
Figure A3 shows that both the total area serviced and the number of active tractors increased rapidly following the introduction of the monitoring technology, reflecting growing demand for digital monitoring among tractor owners. To illustrate the spatial extent of the data, the right panel of Figure 2 maps the geographic distribution of tractor activity across all years, showing that operations cover nearly all agricultural regions of Kenya wherever cropland is present.

3.2 Home Locations

In the absence of consistent information on tractor owner residences, I infer tractor home or base locations from nighttime pings, similar to the literature studying human mobility with smartphone app data. In particular, home locations are defined as the ward (Kenya's

[13] Own calculations based on the cropland mask Digital Earth Africa, where I classify pixels with a probability greater than 70% as cropland. Note that this is a lower bound since a considerable share of mechanized area in the GPS data is not classified as cropland.

Figure 2. Spatial Distribution of Tractor Home Locations and Total Serviced Area



Note: Maps of the number of HT tractors based in each ward (left) and the cumulative serviced area, aggregating all jobs over time, in percent of total surface area (right). County boundaries are shown in black.

smallest administrative unit) with the highest share of pings between 12am and 5am local time.[14]

Home location data is available for 88% of tractors, accounting for 96% of observed tractor activity. The left panel of Figure 2 maps the spatial distribution of inferred tractor home locations, showing that tractors in my sample are dispersed throughout the country. Note that wards vary significantly in size across regions. Given that I use the centroid of home wards as a reference point for my distance calculations, this introduces additional noise. However, it does not affect my main results, which focus on changes over time.

One caveat to inferring home locations from nighttime pings in this way is that they may not necessarily reflect the owner's residence, but rather where the tractor is parked when inactive, such as at the house of the operator or a trusted contact in rural areas. That is why Nairobi, for example, despite being the top location in terms of number of tractors in the Census, is not in my sample, even though many owners likely live there. Nevertheless, Figure A4 shows a positive correlation between the distribution of the number of tractors acrosss counties in my sample using the Census as a comparison.

[14] As I do not have access to the raw data pings myself, the ward level was chosen by HT so as to facilitate the aggregation of pings and maintain the privacy of their customers.

4 Empirical Analysis

Leveraging the temporal and spatial granularity of the tractor GPS data, I provide empirical evidence on how monitoring shapes patterns of tractor mobility and spatial allocation, and affects field-level productivity as measured by satellite data.

4.1 Tractor Mobility and Spatial Reallocation

First, I use within-tractor variation to document how job locations and characteristics change over time, consistent with a reduction in spatial frictions related to monitoring.

Empirical Strategy. Given the way the tractor GPS data are generated, a key empirical challenge is that I do not observe tractors prior to the adoption of the monitoring technology. To address this, I use the first month after adoption as a benchmark for tractors' pre-adoption behavior. This is reasonable assumption because the empirical literature on monitoring suggests that contractual relationships and incentives are slow to adjust, as principals first need to learn how to use the additional information to manage agents more effectively, gradually granting them more autonomy, verifying their actions and realigning incentives over time.[15]

Formally, I implement a within-tractor event study over the post-adoption period, comparing each tractor's activity in the first month after its first recorded job with its behavior in subsequent months. The specification is:

$$D_{it} = \alpha_i + \alpha_{my(t)} + \sum_{k>0} \beta^k e_{it}^k + \varepsilon_{it} \tag{1}$$

where the dependent variable D_{it} is the distance to home averaged across all jobs serviced on day t by tractor i, e_{it}^k are a set of indicators for the k-th month since adoption with $k = 0$ as the reference period; α_i and $\alpha_{my(t)}$ are tractor and month-year fixed effects.

The coefficients of interest, β^k, capture how job distances evolve with monitoring experience, relative to the first month after adoption. If moral hazard between owners and operators limits tractor's spatial reach and the app gradually alleviates this constraint, β^k should increase over time. Tractor fixed effects absorb any time-invariant differences across tractors, so identification relies on within-tractor variation. Month-year fixed ef-

[15] Hubbard (2003) documents lags in returns to monitoring adoption over a period of up to five years. Kelley et al. (2024) find only small contract adjustments within six months of the intervention. Other studies in developing countries attribute slow adoption of monitoring technologies and adaptation to agents' resistance to monitoring (de Rochambeau, 2021; Houeix, 2025) or organizational barriers related to the type of contracts (Atkin et al., 2017).

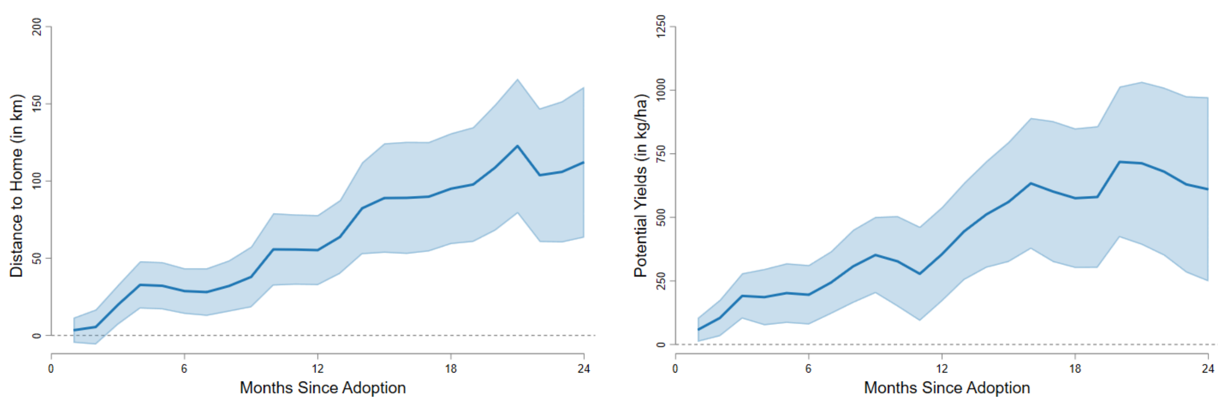
fects control for seasonality and common temporal shocks, which are separately identified from the experience dummies since experience time does not perfectly align with calendar time and tractors enter the sample at different points.[16] While the analysis is not intended to identify causal effects of adoption, it provides a predictive interpretation, where longer use of the app systematically predicts certain behavioral adjustments.

Results. The left panel of Figure 3 plots the estimated coefficients from estimating equation 1 against months since adoption. The coefficients gradually increase by months since adoption, indicating that tractors progressively service jobs further away from their base, as owners gain more experience. On average, tractors operate about 30 km farther from their base location after six months, 55 km after twelve months, and 112 km after twenty-four months, relative to the first month after adoption (with a baseline mean of 70 km). The coefficients at longer durations are estimated less precisely as there are fewer observations.

This pattern is consistent with the interpretation that moral hazard limits the range of operation of tractors and that giving owners the ability to monitor remotely helps alleviate this friction over time. In the model that follows, I use this variation to identify the shifter in the structural parameter capturing monitoring frictions, which underlies the counterfactual simulations.

Figure 3. Within-Tractor Event Study Results

Range of Operation | Returns to Mechanization



Note: Each panel plots estimated coefficients relative to the first month after adoption, with 95% confidence intervals based on robust standard errors, clustered at the tractor level.

The previous result shows that tractors expand into new areas, raising the question

[16]See Appendix B for more details on how the experience dummies are constructed.

of whether this reallocation is also more efficient. I address this by re-estimating equation 1 using a measure of marginal returns to mechanization as the dependent variable. Specifically, I use agro-ecological potential yields from the FAO, computing the difference between the high-input scenario (which includes mechanization) and the low-input scenario (without mechanization). This captures the additional yield a farmer in a particular location could achieve through mechanization, based on exogenous location characteristics (e.g. climate, soil quality). This value is matched to each job by location.[17]

I find that tractors increasingly operate in areas where potential returns to mechanization are higher. After one year, associated average potential yields of serviced jobs are 277 kg/ha higher compared to the first month after adoption (a 6% increase relative to the baseline mean). This pattern is consistent with tractors exploring and learning about the profitability of new locations and suggests that reducing monitoring frictions through the new technology leads to a more efficient spatial allocation of tractors.

Robustness and Additional Results. Although I do not claim causality, the results in this section are central for the subsequent quantitative analysis using the model, so I briefly discuss potential sources of bias and robustness (see Appendix B for additional figures). One concern is selection into adoption. Bias could arise if the timing of adoption is systematically related to future job location choices or to the expansion of tractors' operating range. The direction of such bias, however, is less clear: owners facing severe monitoring problems may adopt earlier and display stronger responses, while more professional owners may adopt early to improve management and would have expanded regardless of monitoring.

Another possible concern is that the observed pattern is driven by entry and exit dynamics, in particular survivorship bias. To limit this, I restrict the analysis to the first two years after adoption, where most tractors are observed. Beyond this period, survivorship bias could arise if only tractors that expand their range remain active (see Appendix B for results including all months for completeness). In addition, I conduct a cohort analysis, estimating the event study separately by entry cohort to check whether early adopters are systematically different, but I find that the gradual increase in distance appears for all adoption cohorts.

I also examine alternative explanations. The observed learning pattern could reflect general learning about the business rather than a reduction in monitoring frictions. However, older tractor models, whose owners had operated the tractor before installing the monitoring device, exhibit similar dynamics, suggesting that the adjustment reflects mon-

[17] See Appendix A for details.

itoring experience rather than generic learning. Another concern is that the increase in distances could be mechanical if tractors are reassigned to distant areas and remain there for extended periods. However, restricting the sample to tractors that 'sleep' at home every night (high-confidence sample), I find similar results. Moreover, the probability of returning to the home ward after working elsewhere remains unchanged, indicating that the effect is not driven by trip chaining.

Finally, I find no evidence that monitoring affects other outcomes related to capacity utilization. In fact, the probability of being active and the number of jobs per day conditional on being active both decline over time, consistent with tractors traveling farther to undertake more productive jobs rather than increasing overall intensity of use.

4.2 Field-Level Productivity

The previous section documented how monitoring shapes the allocation of tractor services, showing that over time tractors expand into new areas with higher potential returns to mechanization. I now complement this analysis by examining whether monitoring translates into improved agricultural outcomes, leveraging the spatial granularity of the data. Specifically, I compare remotely sensed productivity in fields visited by a monitored tractor to that of similar nearby fields in the years before and after the visit.

Empirical Strategy. There are two main empirical challenges in estimating the impact of the monitoring technology on productivity. The first is the lack of disaggregated yield or production data. To address this, I construct a proxy for productivity using satellite-based measures of vegetation dynamics. The second challenge is that there absence of a clear counterfactual, as tractor visits are not randomly assigned and we do not know whether a given field would have been mechanized in the absence of HT. To overcome this, I construct a comparison plot for each eligible job, matching on observable characteristics and pre-treatment trends. The following paragraphs describe these steps in more detail.

To address the lack of yield data, I construct a proxy for productivity from satellite imagery. Specifically, I use the Normalized Difference Vegetation Index (NDVI), which measures vegetation health and density based on surface reflectance, to track crop growth over the season. Unlike previous studies that use NDVI levels directly, I compute the annual difference between the maximum and minimum NDVI values (ΔNDVI_y). Because tractor use at land preparation typically leaves fields bare at the start of the season, this difference provides a more suitable proxy for productivity in the context of mechanization.[18]

To estimate the differential effect of a visit by a monitored tractor, I require a comparison group. Constructing such a group is challenging because satellite data do not directly identify plot boundaries or whether a location has previously been mechanized. Simply comparing treated pixels to all other pixels would produce biased estimates, since treated areas necessarily contain cropland while untreated pixels may include a mix of cropland and non-agricultural land, leading to mechanical differences in vegetation growth.

To address this, I implement a matching procedure to construct a plausible control group of nearby plots that are similar in size, have a comparable probability of being classified as cropland, and exhibit similar pre-treatment vegetation trends. I restrict the analysis to jobs conducted between January and April, corresponding to the typical land-preparation period, and to the first job at a given location to avoid overlap across treatment events. I further restrict the treatment group to jobs whose timing falls within the seasonal window used to compute the NDVI difference.

I construct and match a one-to-one comparison polygon for each treated job as follows. First, I draw a 1 km buffer around the treated job and identify all pixels with a cropland probability within ± 5 percentage points of the treated location. Second, I polygonize contiguous pixels into candidate comparison plots and retain only those whose area is within ± 20 percent of the treated job. Finally, among the remaining candidates, I select the polygon with the most similar pre-treatment trend in ΔNDVI, minimizing the sum of squared differences over the pre-treatment period.

Jobs are sorted chronologically so that when buffers overlap, pixels already selected as comparison plots for one job are iteratively removed and cannot be reused. If no suitable candidate fulfilling all criteria is available, the corresponding job is dropped from the sample. The matching process is computationally intensive, as it requires streaming all available satellite images for each relevant pixel and year to compute the NDVI metrics, and repeating this across multiple years. The final matched sample comprises N = 33,256 job-comparison pairs across Kenya (see Appendix A for a visual example).

Figure 4 plots the average seasonal vegetation growth, measured by ΔNDVI, for treated and control fields in the years before and after the monitored tractor visit, based on the raw data. In the pre-treatment period, both groups follow parallel trends, providing support for the validity of the matching approach. Starting in the year of the tractor visit (event time 0), the NDVI difference diverges sharply for treated fields, suggesting improved vegetation growth consistent with tractor use. The gap between groups appears to persist for several years after the visit, indicating sustained productivity gains.

[18] See Appendix A for more information.

Figure 4. Mean ΔNDVI for Matched Job and Control Fields

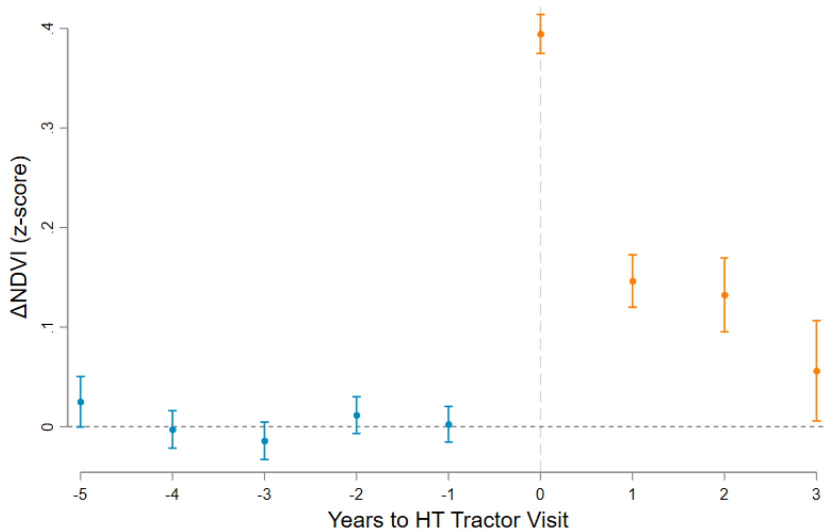
Estimation. I formally estimate the effect of a monitored tractor visit on field producctivity while controlling for potential confounders. Given the matched design, the analysis exploits within-pair variation by comparing each treated plot to its matched counterpart. Specifically, I take the within-pair difference in the outcome and estimate an event-study specification on this pair-wise gap, which isolates the dynamic effect of tractor visits net of time-invariant differences across pairs:

$$\Delta_p(\Delta\text{NDVI}_y) = \mu_p + \lambda_y + \sum_{l=-5}^{3} \beta_l \mathbb{I}\left[y - \text{TractorVisit}_p = l\right] + \varepsilon_{py} \tag{2}$$

where $\Delta_p(\Delta\text{NDVI}_y)$ denotes the outcome difference within pair p between the treated and matched control polygon in year y, μ_p are pair fixed effects, λ_y are year fixed effects.TractorVisit_p indicates the year of the tractor visit. The event study coefficients β^l capture the dynamic treatment effect of tractor visits on the NDVI gap over time. I estimate this specification using the estimator by Callaway and Sant'Anna (2021) to account for staggered treatment timing and heterogeneous treatment effects. Identification relies on parallel trends and no anticipation assumptions, which are likely to hold due to the matching procedure. SUTVA is also likely to hold since eligible pairs are spread out over space, reducing the likelihood of spillovers.

It is important to note that control fields may or may not have been mechanized, reflecting the average level of mechanization in the broader population. The estimated difference therefore represents a lower bound on the productivity effect of mechanization and can be interpreted as the additional gains associated with monitored tractor services relative to the prevailing average.

Figure 5. Effect on Field-Level Productivity



Results. Figure 5 shows the estimated effect of tractor visits on the within-pair outcome over time. Consistent with the raw data patterns, there is a sharp increase in the year of the tractor visit of about 0.4 standard deviations-nearly three times the pre-treatment mean of the pair-wise difference. This indicates that a visit by a monitored tractor had a clear differential effect compared to similar plots in the same location. The effect also exhibits some persistence, suggesting that plots benefited from mechanization in terms of productivity in subsequent years.

Decomposing the NDVI measure into its components (Figure B10) reveals that the immediate jump is driven mainly by a decrease in NDVI at the start of the season, consistent with newly induced mechanization and more effective land preparation, while the persistence reflects a sustained increase in the maximum NDVI value, indicating productivity gains from improved crop growth following mechanization.

Taken together, the empirical findings point to meaningful efficiency gains from reducing monitoring frictions in tractor rental markets via digital monitoring. Tractors expand their range of operation over time, reallocating toward areas with higher potential returns, and their visits are associated with measurable improvements in field-level vegetation growth. However, it remains empirically challenging to translate these effects into aggregate output impacts. The relationship between NDVI and actual crop yields is non-linear and context-specific, and the empirical analysis does not capture general equilibrium responses. Therefore, to quantify the broader implications of reducing monitoring frictions, I develop a spatial model that links these micro-level behavioral responses to aggregate outcomes in the next section.

5 Quantitative Spatial Model of Tractor Location Choice

In this section, I develop a quantitative spatial model to quantify the aggregate gains from monitoring, taking into account equilibrium effects of reallocating scarce capital across space.

The framework shares the gravity structure of standard models used to study the flow of goods in trade or the movement of people in migration and, more recently, commuting within cities in the urban economics literature (Ahlfeldt et al., 2015; Tsivanidis, 2023). However, it departs from these applications in important ways, adapting the model to capture the mobility of capital in the form of tractors and extending it to incorporate monitoring costs that are spatial in nature. Instead of workers choosing home and work locations, tractor owners choose where to send their tractors given their base location and moving costs that depend on the monitoring technology available to them.

5.1 Setup

Consider an economy with a discrete set of locations $j = 1, ..., J$. A measure Ω of tractor owners holds a fixed stock of tractors $\bar{T}$, with an initial distribution across home locations $T_{-1,h}$ ($h \in J$).[19] In each location j, there is a representative farmer who produces a homogeneous, freely traded good (with price normalized to $P = 1$) using land and tractor services as inputs.

5.2 Tractors

A tractor owner ω chooses where to send the tractor given their home location h, maximizing the value of working in j:

$$\max_{j \in N} V_{j|h}(\omega) = \frac{r_j}{c_{hj}} z_j(\omega) \tag{3}$$

[19] A natural extension would be to endogenize the stock of tractors, allowing entry to respond to the profitability of renting them out.

where r_j denotes the return earned by tractors in location j, c_{hj} is the cost of moving between h and j, and $z_j(\omega)$ is an idiosyncratic productivity draw specific to location j.

Moving costs c_{hj} negatively scale payoffs and capture both direct costs related to moving the tractor through space, such as fuel or maintenance, and indirect costs related to monitoring (arising from the owner-operator moral hazard discussed above). They take the form

$$c_{hj} = D_{hj}^{\kappa} \tag{4}$$

where D_{hj} is the bilateral distance between h and j, and κ is the elasticity of moving costs with respect to distance.[20] This elasticity governs how quickly operating costs rise with distance and is the key parameter of interest.

Importantly, κ can be decomposed into two components, $\kappa = \kappa^T + \kappa^M$, where κ^T captures sensitivity of distance-based costs purely due to transport (physical), and κ^M captures the additional sensitivity arising from imperfect monitoring (behavioral). Better monitoring, such as through the use of the new digital monitoring technology, reduces κ^M, since for the same distance, owners perceive the cost of operating the tractor as lower.

The idiosyncratic term $z_j(\omega)$ captures heterogeneity in tractor owners' productivity across locations, reflecting factors such as previous customer relationships, social networks, or prior knowledge about local demand. These productivity draws are assumed to be independent and identically distributed Fréchet, with cumulative distribution function $F(z) = \exp(-z^{-\theta})$. The shape parameter θ governs the degree of heterogeneity across owners: lower values imply greater dispersion, meaning that some owners have particularly strong connections or comparative advantages in certain areas.

Under the Fréchet assumption, the probability that a tractor based in location h operates in location j is given by (Eaton and Kortum, 2002)

$$\pi_{j|h} \equiv Pr\left[\frac{r_j}{c_{hj}} z_j > \max_{j' \neq j} \frac{r_{j'}}{c_{hj'}} z_{j'}\right] = \frac{(r_j/c_{hj})^{\theta}}{\sum_s (r_s/c_{hs})^{\theta}} = \frac{(r_j/c_{hj})^{\theta}}{\Phi_h} \tag{5}$$

where the denominator, Φ_h, is constant across destinations and summarizes the overall market access that tractors in h have to profitable locations.

Taking logs of this conditional probability and subsituting for c_{hj} yields the gravity

[20] In standard commuting models, moving costs are typically expressed as an exponential function of travel time, where κ takes the interpretation of a semi-elasticity of costs with respect to time or distance. Instead, I opt for the power function formulation, which offers a better empirical fit in this context. When calibrating the model, I impose the numerical approximation $D_{hj} \geq 1$ to ensure the logarithm of distance is defined.

equation:

$$\log \pi_{j|h} = \theta \log r_j - \theta\kappa \log D_{hj} - \log \Phi_h \tag{6}$$

Intuitively, tractors based in h are more likely to operate in j when returns are higher and when moving costs are lower. The strength of these responses depends on the shape parameter θ, because a higher θ implies less idiosyncratic variation across destinations and thus greater sensitivity of location choices to returns and distance costs, while a lower θ implies the opposite.

By the law of large numbers, $\pi_{j|h}$ also represents the share of tractors from h that operate in j. The total supply of tractors to any destination is therefore obtained by aggregating the number of tractors that move there over all origins:

$$T_j = \sum_h \pi_{j|h} T_{-1,h} = \sum_h \frac{(p_j/c_{hj})^\theta}{\Phi_h} T_{-1,h} \tag{7}$$

5.3 Farmers

Instead of modeling discrete tractor adoption decisions at the farm level, I model production at the aggregate level using a representative farmer. Tractor use T_j represents a continuous measure of the average mechanization levels in location j, corresponding to the share of farmers who use a tractor. This simplifies the analysis of spatial equilibrium while preserving the key economic mechanisms.[21]

The representative farmer in location j produces a homogenous good using a Cobb-Douglas technology with decreasing returns to tractor services ($\gamma < 1$):

$$Y_j = A_j \bar{H}_j^{1-\gamma} T_j^\gamma \tag{8}$$

where A_j is exogenous productivity of location j, and $\bar{H}_j$ is a fixed land endowment. Farmers choose the amount of tractor services T_j to maximize profits, taking land as given.

The first-order condition equates the value of the marginal product of tractor services to the return r_j:

$$r_j = \gamma A_j \bar{H}_j^{1-\gamma} \left(\frac{T_j}{\bar{H}_j}\right)^{\gamma-1} \tag{9}$$

[21] An extension of the model could incorporate the extensive margin of adoption more explicitly. In this paper, I abstract from this dimension in order to focus on the misallocation of tractors across space conditional on adoption.

This expression implies a downward-sloping demand for tractors. In equilibrium, returns r_j depend on exogenous productivity A_j, which captures how productive tractors are in that location (e.g. soil quality or seasonality), and on local density of tractors $T_j/\bar{H}_j$, which limits how many tractors can be profitably employed.

5.4 Equilibrium

Given exogenous and geographic features $(A_j, \bar{H}_j, D_{hj}, T_{-1,h})$, model parameters (θ, κ, γ) and the total stock of tractors $\bar{T}$, an equilibrium is a vector of endogenous objects (T_j, r_j) such that:

1. *Tractor markets clear*: For all locations j, the supply of tractors $T_j = \sum_h \pi_{j|h} T_{-1,h}$ is consistent with farmers' demand $r_j = \gamma A_j (T_j/\bar{H}_j)^{\gamma-1}$;

2. *Total tractor supply is fixed*: $\bar{T} = \sum_j T_j = \sum_h T_{-1,h}$.

6 Model Estimation

In this section, I bring the model to the data. I estimate key parameters from observed tractor flows, calibrate the demand side using my farmer survey, and invert the model to recover location-specific productivities that rationalize the spatial distribution of tractor utilization in the data.

6.1 Tractor Location Choice Parameters

Using the model's gravity structure, I jointly estimate the elasticity of tractor flows θ and the average distance elasticity of moving costs $\bar{\kappa}$. I introduce a new approach to separately identify θ, leveraging exogenous variation in potential yields as a shifter to returns across locations. The parameter $\bar{\kappa}$ captures an average moving cost elasticity, including any monitoring effects, and is used for model inversion; in the next subsection I allow κ to vary with monitoring experience for the counterfactual analysis.

The model yields a simple gravity equation (6) that identifies the parameter cluster $\beta = -\theta\bar{\kappa}$ from the sensitivity of tractor flows to distance between any two locations, controlling for origin (α_h) and destination (α_j) fixed effects:

$$\log \pi_{j|h} = \alpha_j + \beta \log D_{j|h} + \alpha_h + \eta_{j|h} \tag{10}$$

where $\pi_{j|h}$ is the bilateral tractor flow, measured as the number of jobs serviced in location j by tractors based in h, pooled across all tractors and time, and $D_{j|h}$ is the straight-line distance between respective ward centroids.

Following the literature, I estimate this relationship via Pseudo Poisson Maximum Likelihood (PPML) with high-dimensional fixed effects to account for zeros in the data (Correia et al., 2020). The results are shown in Column (1) of Table 1. The coefficient on distance implies that a 1% increase in distance reduces tractor flows by about 1.3%. For example, increasing the distance between two wards from 10 km to 20 km (a 100% increase) is associated with roughly a 75% decline in the number of jobs serviced between them, consistent with the base location being the main determinant of tractor allocation.

Table 1. Pooled Gravity Estimation Results

	(1) PPML	(2) OLS
Log(Distance)	-1.295***	
	(0.035)	
Log(Potential Yields)		0.684***
		(0.204)
Terrain Ruggedness Index		-0.218***
		(0.015)
Relative Wealth Index		-2.615***
		(0.270)
Log(Population)		0.553***
		(0.162)
Log(Surface Area)		0.305***
		(0.083)
Obs.	401,388	986
Pseudo R^2	0.69	
R^2		0.40

Column (1) reports estimates of the gravity equation (10) using PPML with two-way fixed effects. The outcome variable is HT tractor flow, $\pi_{j|h}$, measured as the number of jobs serviced in j by tractors based in h, aggregated across tractors and time. Two-way clustered standard errors at the origin and destination level are reported in parentheses. Column (2) reports results from the OLS regression of equation (11) at the ward level, where the dependent variable is the destination fixed effect estimated in (1). See Section A for details on the data used. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

I propose a new approach to identify θ, which then allows me to back out $\bar{\kappa}$. The existing literature, including Ahlfeldt et al. (2015), typically calibrates θ by relating the dispersion of observed prices or wages across locations to the variance of the Fréchet distribution, since its dispersion parameter governs the elasticity of trade or commuting flows in the model. However, this strategy is very indirect and may conflate structural heterogeneity with idiosyncratic noise, measurement error, or other forces outside the model. Instead, I directly estimate θ from the behavioral response of tractor flows to an exogenous shift in returns (since in the model θ is the coefficient on $\log r_j$). The idea is that tractor flows should respond to an observed shock in the same way they would respond to unobserved productivity draws, providing a cleaner and more direct identification.

I implement this strategy by recovering the estimated destination fixed effects $\hat{\alpha}_j$ from equation 10, which capture residual tractor flows to a destination j beyond what can be explained by distance, and regress them on potential yields $\tilde{r}_j$, which serve as an exogenous shifter of returns:

$$\hat{\alpha}_j = \alpha_0 + \theta \log \tilde{r}_j + \delta' X_j + \varepsilon_j \tag{11}$$

where X_j are controls for other observed location characteristics (terrain ruggedness, wealth, population and surface area) that may influence tractor utilization in j and are included to avoid omitted-variable bias.

Column (2) of Table 1 reports the results, where the coefficient on potential yields provides $\hat{\theta} = 0.684$. This relatively low estimate of θ implies substantial heterogeneity across tractor owners in their location-specific productivity draws and limited responsiveness of tractor flows to spatial differences in returns, consistent with the idea that some owners are better connected or more familiar with certain areas.[22] Then, $\bar{\kappa}$ can be computed as $-\hat{\beta}/\hat{\theta} = 1.295/0.684 = 1.893$.

6.2 Moving Cost Elasticities by Monitoring Experience

Next, I estimate how the elasticity of moving costs, κ, varies with monitoring experience. The estimation exploits the same empirical variation arising from the gradual learning dynamics documented in Section 4.1, and the resulting estimates serve as inputs for the comparative statics exercises in Section 7.

I estimate a gravity regression at the tractor level, allowing the distance coefficient to depend linearly on months since adoption:

$$\log \pi_{j|h(i),m} = \alpha_{jmy} + \beta_0 \log D_{j|h(i)} + \beta_1 \tau_{im} \times \log D_{j|h(i)} + \alpha_i + \eta_{j|h(i),m} \tag{12}$$

where $\pi_{j|h(i),m}$ denotes the number of jobs serviced in destination j by tractor i (based in $h(i)$) during calendar month m, and τ_{im} measures cumulative monitoring experience in months since adoption. Tractor fixed effects α_i replace the origin fixed effects in the

[22]The dependent variable in equation (11) is an estimated destination fixed effect. Measurement error in $\hat{\alpha}_j$ could attenuate $\hat{\theta}$ if precision varied systematically with potential yields. In practice, potential yields are not strongly correlated with the number of jobs in the raw data (although they do of course affect residual bilateral flows conditioning on controls), suggesting this concern is limited.

pooled gravity, absorbing time-invariant heterogeneity across tractors, and destination-month-year fixed effects α_{jmy} capture seasonality.[23]

The coefficient β_0 identifies the baseline sensitivity of individual tractor flows with respect to distance in the first month of adoption, and β_1 captures how this elasticity changes with each additional month of experience. Conceptually, this specification provides a model-based analogue to the dynamic effects after monitoring adoption documented in Section 4.1.

Table 2 reports the results of estimating equation 12 using PPML. The coefficient on distance is larger in magnitude than the pooled estimate, consistent with the pooled regression averaging across tractors at different stages of adoption, which includes the effect of monitoring. The interaction term, $\hat{\beta}_1 = 0.00233$, indicates that this elasticity becomes slightly less negative with each additional month of monitoring experience: as tractors accumulate experience using the technology, their effective sensitivity to distance declines. Both effects are highly significant at the 5% level, consistent with monitoring lowering perceived moving costs over time, allowing tractors to operate profitably over longer distances.

Table 2. Tractor-Level Gravity Estimation Results

	(1)
Log(Distance)	-1.32989***
	(0.03201)
Log(Distance) $\times\ \tau$	0.00233***
	(0.00090)
Obs.	10,739,033
Pseudo R^2	0.54

This table report estimates of the tractor-level gravity equation (12) using PPML with two-way fixed effects. The outcome variable $\pi_{j|h(i),m}$ are flows at the tractor $\times$ month level. Two-way clustered standard errors at the tractor and destination level are reported in parentheses. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Finally, I obtain the structural elasticities by dividing the estimated coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ by $\hat{\theta}$ from before, such that $\hat{\kappa} = 1.944$ and $d\hat{\kappa}/d\tau = -0.0034$. Here, $\hat{\kappa}$ represents the overall baseline distance elasticity of moving costs, and $\frac{d\hat{\kappa}}{d\tau}$ captures how this elasticity declines with monitoring experience, corresponding to a reduction in κ^M in the model.

[23] Because the data is monthly, calendar time and experienced time are perfectly collinear, so I cannot separately include month-year fixed effects as before. Since origins don't vary over time within tractor, I include destination-month-year fixed effects instead.

6.3 Demand

The tractor rental cost share γ is calibrated using data from my farmer survey. Specifically, I compute γ as the average share of tractor rental expenditures in total production costs (including expenditures on seeds, fertilizer, pesticides, hired labor, and transportation) across all farmers.[24] Non-mechanized farmers are included to reflect the model assumption that some producers do not use tractors. The resulting calibration is $\gamma = 0.14$, indicating that tractor rental costs account for about 14% of total production costs on average.

6.4 Model Inversion

Given the estimated parameters, I invert the model to recover the equilibrium returns and implied productivities that rationalize the observed allocation of tractors. Specifically, I use the tractor supply equation (7) together with the pooled estimate of $\bar{\kappa}$ from Section 6.1 and data on the number of tractor units at origin $T_{-1,h}$ and destination T_j (aggregated from the pooled flow data) to solve for the equilibrium returns r_j via a fixed-point algorithm. I then back out the implied productivities across locations, A_j, from the tractor demand schedule, using ward-level cropland area as input for $\bar{H}_j$ (see Section A for details).

7 Quantifying the Gains from Monitoring

In this section, I use the estimated model to quantify the aggregate gains from reducing monitoring frictions by simulating counterfactual equilibria under different values of the moving cost elasticity, corresponding to different stages of monitoring experience. I examine how these changes affect the spatial allocation of tractors and overall productivity, first within the HT sample and then in the aggregate economy, and compare the costs of an equivalent fuel subsidy calibrated to deliver the same efficiency gains.

7.1 Gains from Monitoring among HT Tractors

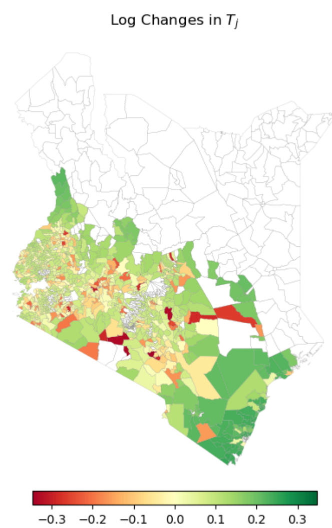
In this section, I consider an economy composed of only HT tractors (within-sample) and simulate two equilibria using the estimates from Section 6.2 as inputs: (i) a baseline scenario removing monitoring benefits, setting $\kappa = \hat{\kappa} = 1.94$, where tractors behave as if they had just adopted the technology; and (ii) a situation in which all tractors behave as if they had accumulated five years of monitoring experience, corresponding to $\kappa' = \hat{\kappa} + 60 \times d\hat{\kappa}/d\tau = 1.74$. Both equilibria are solved under the same initial distribution of

[24]Sampling weights are applied to account for the oversampling of farmers that have received a HT visit in the survey.

tractors across origins and reallocating the same total number of jobs per tractor, such that differences in outcomes arise solely from changes in monitoring.[25]

Figure 6 shows the predicted log change in tractor supply across locations, between the baseline and the improved monitoring scenario. Tractor utilization shifts away from central areas and toward more remote regions. Comparing the correlation between the distribution of T_j and T_h, as well as between T_j and A_j, indicates that this reallocation is driven by tractors moving away from their bases and toward more productive locations.

Figure 6. Predicted Change in Tractor Supply Across Locations



To quantify the total efficiency gains from monitoring, I calculate the percentage change in aggregate output across locations as well as the percentage change in the standard deviation of the marginal products of tractor services, which serves as a measure of spatial misallocation. Relative to the baseline scenario, improved monitoring leads to a 2.18% increase in total output, driven by a 15.9% reduction in the dispersion of marginal products–that is, a more efficient spatial allocation of tractors across locations (see Table 3). These results highlight the substantial productivity benefits that arise from reducing monitoring frictions in tractor rental markets.

Table 3. Counterfactual Results: Baseline vs. 5 Years of Monitoring

	$\%\Delta \sum_j Y_j$	$\%\Delta\sigma(\partial Y_j/\partial T_j)$
Within-Sample	2.18	-15.9
Aggregate	1.30	-9.0

[25]This exercise differs from the empirical patterns in the data, where tractors adopt monitoring at different times and are observed for varying durations. The counterfactual instead holds the tractor stock fixed and imposes a common monitoring regime, isolating the equilibrium reallocation effects.

7.2 Rolling Out Monitoring to All Tractors

I next extend the analysis to the aggregate economy that includes all tractors in Kenya, not only those in the HT sample. This exercise is equivalent to assessing what would happen if digital monitoring were rolled out to all tractors nationwide. While I do not observe where non-HT tractors operate, the 2019 Population Census provides information on the number of tractor-owning households at the county level. I use these data, together with the model estimated from HT flows, to predict the equilibrium allocation of tractors under κ and κ' as above, thereby demonstrating the model's applicability to out-of-sample policy analysis.

The Census reports the number of tractor-owning households in a 10% sample, which I scale by a factor of ten to approximate total counts. I then weight each tractor by the average number of jobs in the HT data to express utilization in comparable units. To obtain a ward-level distribution, I disaggregate the county totals across wards within each county as follows. For counties with HT home locations, I allocate Census tractors proportionally to each ward's share of origin jobs ithe county total in the HT data. For the six counties without any HT tractors (Mombasa, Mandera, Wajir, Isiolo, Murang'a, and Elgeyo-Marakwet), I estimate a ward-level regression of HT origin jobs on the income (measured by the Relative Wealth Index) and cropland area, including county fixed effects, using data from counties with HT coverage. I then use the estimated coefficients to predict ward-level job counts in the non-HT counties, normalize these predictions within each county, and scale them to match the county's Census total. This yields a new, aggregate initial distribution of tractors at the ward level, $T^{agg}_{-1,h}$, that is consistent with observed home locations.[26]

Repeating the same two counterfactuals using this aggregate initial distribution, I find that rolling out monitoring to all tractors would raise aggregate output by 1.3% and reduce the dispersion of marginal products by 9% (see Table 3). This pattern is qualitatively similar to the within-HT counterfactual, though the magnitudes are smaller, which is expected given that the aggregate scenario incorporates the full tractor distribution, including units outside the HT sample that may already be located in more productive areas.[27]

[26] I do not combine Census and HT tractor counts to avoid double counting, as tractors appearing in HT are a partial subset of the Census totals. The counterfactual therefore relies solely on Census daata, where I used the HT data to inform the within-county disaggregation.

[27] A remaining caveat is that destinations where no HT tractors are ever observed do not enter the model, since I cannot recover local productivities A_j for these locations in the model inversion step. As a result, the aggregate counterfactual may understate total gains if monitoring would enable tractors to profitably expand into areas that are unobserved in the HT data.

7.3 Comparison with Fuel Subsidy

Finally, I compare the monitoring-based efficiency gains to an alternative policy the government could implement to encourage tractors to operate farther from their home bases, thereby improving access to mechanization for remote farmers. Specifically, I consider an equivalent fuel subsidy. Unlike digital monitoring, which reduces the distance elasticity of moving costs κ, this policy lowers the overall level of transport costs c_{hj} directly, while holding κ fixed at its baseline value $\hat{\kappa}$.[28] I first calibrate the required reduction in effective distance costs to match the efficiency gains from the five-year monitoring rollout, and then approximate the total cost of such a subsidy to compare it with the cost of digital monitoring using a back-of-the-envelope calculation.

Using the model, I compute the proportional reduction x per unit of distance needed to replicate the monitoring gains in the aggregate economy, such that $c'_{hj} = (D_{hj}(1-x))^{\hat{\kappa}}$. The required reduction is large: distances would need to fall by $x = 0.71$ per km across the board. To translate this into a fuel subsidy, I consider the share of per-kilometer variable transport costs attributable to fuel.[29] Under these assumptions, the subsidy would need to cover essentially 100% of per-kilometer fuel expenditures, implying a cost of 1.05 USD per km multiplied by the total kilometers travelled by all tractors, $\sum(\pi'_{j|h}T_{-1,h})\, D_{hj}$. Using model-predicted tractor flows, this corresponds to 9.33 billion USD.

In contrast, equipping all 70,930 Census tractors with a GPS device (125 USD) and a five-year subscription (60 USD per year) costs only 30.1 million USD - less than one percent of the equivalent fuel-subsidy cost. This comparison highlights that digital monitoring provides a highly cost-effective means of improving the spatial efficiency of tractor rental markets.

8 Conclusion

This paper offers a new perspective on barriers to agricultural mechanization, showing how *supply*-side monitoring frictions constrain the mobility of productive capital and, as a result, limit the allocation of tractor services to areas where they are most needed. Specifically, the paper evaluates a reduction in distance-based monitoring cost through the introduction and expansion of a new GPS tracking app for tractors in Kenya, leveraging the

[28]This comparative statics exercise is similar to reducing travel times or transport infrastructure improvements in the urban economics literature.

[29]Assuming fuel consumption of 0.8 L/km and a diesel price of 170 KES/L, fuel costs amount to 136 KES/km. Adding standard values for lubrication (16 KES/km), maintenance (34 KES/km), and operator wages (8 KES/km) implies that fuel constitutes roughly 70% of variable movement costs. Achieving a 71% reduction in effective distance therefore requires subsidizing nearly all of the per-kilometer fuel cost.

universe of GPS tractor activity records from its adopters. Combining detailed empirical evidence with a quantitative spatial model, the analysis demonstrates how digital monitoring facilitates a more efficient spatial allocation of tractors and generates substantial (aggregate) productivity gains in agriculture.

The findings also highlight the broader potential of digital technologies as a scalable, cost-effective tool to alleviate information frictions and expand access to mechanization. While the empirical context is Kenya, the mechanism is likely relevant across both countries and sectors. In agriculture, many Sub-Saharan African economies share Kenya's low mechanization levels and institutional features, but even more developed settings rely on rental markets and migratory service provision for expensive machinery, such as combine harvesters in Spain. Beyond agriculture, sectors that rely on the delegated operation of mobile capital—such as trucking, ride-hailing, or public transportation—face analogous monitoring frictions that can constrain spatial mobility and efficiency. Reducing these information asymmetries can therefore generate meaningful productivity gains well beyond the agricultural sector.

The analysis also has limitations that open avenues for future research. Empirically, productivity is proxied through vegetation dynamics rather than plot-level yields. Improved yield data would allow a tighter mapping between NDVI changes, farm profits, and welfare.[30] I also plan to examine how the monitoring-induced expansion of mechanization affects local economic development and structural transformation by linking the tractor data to nationally representative surveys such as the DHS, although this poses empirical challenges. On the modeling side, the framework holds the tractor stock fixed and focuses on the intensive spatial margin; allowing for endogenous entry and capital deepening could amplify the aggregate effects and help assess implications for market structure and competition among service providers. Finally, the model abstracts from destination-side coordination and matching frictions, which are currently absorbed into idiosyncratic productivity draws; empirically exploring these frictions and quantifying their role remains a promising direction for future work.

Taken together, the results suggest that digital technology can yield potentially large efficiency gains by easing agency frictions in markets for mobile capital. Understanding how such organizational improvements interact with spatial frictions remains an important area for future research on productivity and structural transformation in developing economies.

[30] In related work, I am extending the county-level yield predictions developed in Bohra et al. (2025) to the field level.

References

Aggarwal, S., B. Giera, D. Jeong, J. Robinson, and A. Spearot (2024). Market access, trade costs, and technology adoption: Evidence from northern tanzania. *Review of Economics and Statistics 106*(6), 1511–1528.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Atkin, D., A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen (2017). Organizational barriers to technology adoption: Evidence from soccer-ball producers in pakistan. *The Quarterly Journal of Economics 132*(3), 1101–1164.

Baker, G. P. and T. N. Hubbard (2003). Make versus buy in trucking: Asset ownership, job design, and information. *American Economic Review 93*(3), 551–572.

Bassi, V., R. Muoio, T. Porzio, R. Sen, and E. Tugume (2022). Achieving scale collectively. *Econometrica 90*(6), 2937–2978.

Belton, B., M. T. Win, X. Zhang, and M. Filipski (2021). The rapid rise of agricultural mechanization in myanmar. *Food Policy 101*, 102095.

Blanchard, P., D. Gollin, M. Kirchberger, and M. Peters (2025). High-frequency human mobility in three african countries. Working Paper.

Bohra, A., S. Nottmeyer, C. Ren, S. Chen, and Y. Ma (2025). Advancing corn yield mapping in kenya through transfer learning. *Remote Sensing 17*(10), 1717.

Bryan, G., K. Frye, and M. Morten (2025). Chapter 10 - spatial economics for low- and middle-income countries. In D. Donaldson and S. J. Redding (Eds.), *Handbook of Regional and Urban Economics*, Volume 6 of *Handbook of Regional and Urban Economics*, pp. 653–714. Elsevier.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230.

Caunedo, J. and N. Kala (2021). Mechanizing agriculture. Technical report, National Bureau of Economic Research.

Caunedo, J., N. Kala, and H. Zhang (2022). Economies of density and congestion in the sharing economy. Technical report, National Bureau of Economic Research.

Correia, S., P. Guimarães, and T. Zylkin (2020). Fast poisson estimation with high-dimensional fixed effects. *The Stata Journal 20*(1), 95–115.

de Rochambeau, G. (2021). Monitoring and intrinsic motivation: Evidence from liberia's trucking firms. Working Paper.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Fischer, G., F. O. Nachtergaele, H. Van Velthuizen, F. Chiozza, G. Franceschini, M. Henry, D. Muchoney, and S. Tramberend (2021). *Global agro-ecological zones v4–model documentation*. Food & Agriculture Org.

Foster, A. D. and M. R. Rosenzweig (2022). Are there too many farms in the world? labor market transaction costs, machine capacities, and optimal farm size. *Journal of Political Economy 130*(3), 636–680.

Houeix, D. (2025). Asymmetric information and digital technology adoption: Evidence from senegal. Working Paper.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics 124*(4), 1403–1448.

Hubbard, T. N. (2000). The demand for monitoring technologies: the case of trucking. *The Quarterly Journal of Economics 115*(2), 533–560.

Hubbard, T. N. (2003). Information, decisions, and productivity: On-board computers and capacity utilization in trucking. *American Economic Review 93*(4), 1328–1353.

Kaumbutho, P. and H. Takeshima (2023). Mechanization of agricultural production in kenya: Current state and future outlook. In C. Breisinger, M. Keenan, J. Mbuthia, and J. Njuki (Eds.), *Food Systems Transformation in Kenya: Lessons from the Past and Policy Options for the Future*, Chapter 9, pp. 231–260. Washington, DC: International Food Policy Research Institute (IFPRI).

Kelley, E. M., G. Lane, and D. Schönholzer (2024). Monitoring in small firms: Experimental evidence from kenyan public transit. *American Economic Review 114*(10), 3119–3160.

Kreindler, G., J. T. Dean, and O. Mbonu (2025, May). The effect of exposure: Evidence from spatial choices in nairobi. Working Paper.

Kreindler, G. E. and Y. Miyauchi (2023). Measuring commuting and economic activity inside cities with cell phone records. *Review of Economics and Statistics 105*(4), 899–909.

Manuelli, R. E. and A. Seshadri (2014). Frictionless technology diffusion: The case of tractors. *American Economic Review 104*(4), 1368–1391.

Mrema, G., J. Kienzle, and J. Mpagalile (2018). Current status and future prospects of agricultural mechanization in sub-saharan africa (ssa). *Agricultural Mechanization in Asia, Africa and Latin America 49*(2), 13–30.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics 94*(1), 20–36.

Olmstead, A. L. and P. W. Rhode (2001). Reshaping the landscape: the impact and diffusion of the tractor in american agriculture, 1910–1960. *The Journal of Economic History 61*(3), 663–698.

Restuccia, D. and R. Rogerson (2017). The causes and costs of misallocation. *Journal of Economic Perspectives 31*(3), 151–174.

Sims, B., M. Hilmi, and J. Kienzle (2016). Agricultural mechanization: a key input for sub-saharan africa smallholders.

Storeygard, A. (2025). Transport in low-and middle-income countries. Technical report, National Bureau of Economic Research.

Suri, T. and C. Udry (2022). Agricultural technology in africa. *Journal of Economic Perspectives 36*(1), 33–56.

Tsivanidis, N. (2023). Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio. Working Paper. Conditionally accepted at American Economic Review.

Walker, M. W., N. Shah, E. Miguel, D. Egger, F. S. Soliman, and T. Graff (2024). Slack and economic development. *NBER Working Paper w33055*, –. Working Paper.

Yang, J., Z. Huang, X. Zhang, and T. Reardon (2013). The rapid rise of cross-regional agricultural mechanization services in china. *American Journal of Agricultural Economics 95*(5), 1245–1251.

A Data Appendix

A.1 Own Farmer Survey

I conducted my own survey of farmers, including both customers of HT and non-customers, across 4 sub-locations in Kakamega County, using spatial sampling within a 2km radius around a market place. I collected agriculural production data, including plot-level tractor and oxplough use as well as crop yields, rental transaction data and GPS plot boundaries of 381 farmers and 429 plots. This data is used for qualitative information, validation and calibration of the model.

A.2 Potential Yields

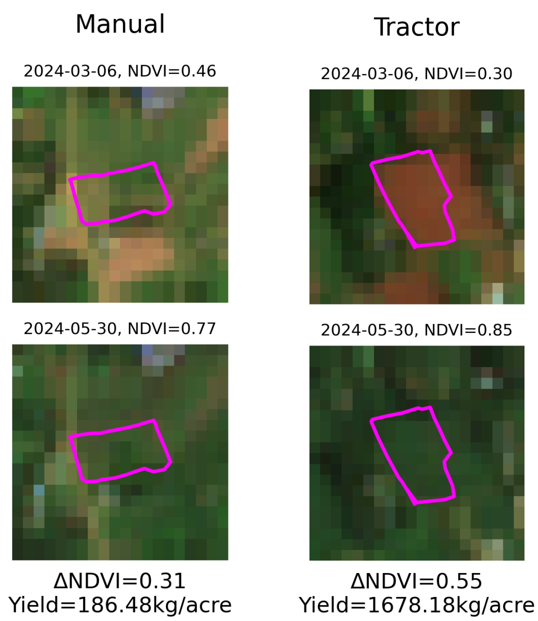
My measure of mechanization gains comes from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) v4 database. This database provides estimates of potential crop yields at 9km resolution, calculated by incorporating local soil and weather characteristics into an agronomic model that predicts maximum attainable yields for each crop in a given area. I focus on maize, which is the most widely grown crop in Kenya. Further, I select the data series applying the average climate over the period 1981-2010 and assumes rain-fed conditions. In addition, the database reports potential yields for different production technologies, a high and a low input scenario. Under high input conditions, production is assumed to be fully mechanized with low labor intensity, whereas low inputs means traditional, labor-intensive farming techniques are used. For every cell, I calculate the difference in potential yields between the two scenarios, which gives me a measure of how much more productive farmland in each location would be if it were fully mechanized, solely based on exogenous location characteristics. I then match this information to job polygons performing a spatial join.

The description of each technology in the FAO-GAEZ documentation is as follows (Fischer et al., 2021). Low-level inputs/traditional management: *"Under the low input, traditional management assumption, the farming system is largely subsistence based and not necessarily market oriented. Production is based on the use of traditional cultivars (if improved cultivars are used, they are treated in the same way as local cultivars), labor intensive techniques, and no application of nutrients, no use of chemicals for pest and disease control and minimum conservation measures."* High-level inputs/advanced management: *"Under the high input, advanced management assumption, the farming system is mainly market oriented. Commercial production is a management objective. Production is based on improved high yielding varieties, is fully mechanized with low labor intensity and uses optimum applications of nutrients and chemical pest,* *disease and weed control."*

A.3 Satellite Data

I use Sentinel-2, Level 2A surface reflectance data from the EU Copernicus mission at 10m resolution to calculate the Normalized Difference Vegetation Index from the relevant bands $NDVI = NIR - RED/NIR + RED$ for every scene from 2017 to 2024, masking for clouds. Start and peak of the growing season are determined by the first increase of the greening curve and the maximum in the NDVI time series, respectively, at the pixel-level. Values at the plot level are the average over all intersecting pixels.

Figure A1. Illustration of ΔNDVI$_t$



Note: Two example fields in Butere, Kakamega County, with information on power source at land preparation and crop yields from farmer survey.

A.4 Matched Sample Construction

Figure A2. Illustration of Job Polygon and Matched Comparison Polygon

A.5 Other Data Sources

Cropland data: I use Digital Earth Africa's 2019 Cropland product at 10m resolution to identify pixels that likely contain cropland. The mask is the output of a machine learning algorithm that was trained on data from Africa specifically and is more accurate and recent than other widely available global land use classifications. I identify a pixel as cropland if the predicted cropland probability exceeds 70%.

Population Census: I use the Kenya Population and Housing Census (10% sample) from 2019 to assess household tractor ownership at the county level.

Spatial units: There are 47 counties and 1450 wards (with a median surface area of 77km^2). I use shapefiles for administrative boundaries from the Humanitarian Data Exchange (HDX).

Terrain Ruggedness Index: To capture local topographic variability, I construct a Terrain Ruggedness Index (TRI) using elevation data from the Copernicus Digital Elevation Model (GLO-90) at 90 m spatial resolution. For each pixel, the TRI is defined as the mean absolute difference between the focal pixel's elevation and the elevations of its eight adjacent neighbors in a 3×3 moving window (Nunn and Puga, 2012):

$$TRI(i,j) = \frac{1}{8} \sum_{(m,n)\in\mathcal{N}_{i,j}} |\, z_{m,n} - z_{i,j} \,| \tag{13}$$

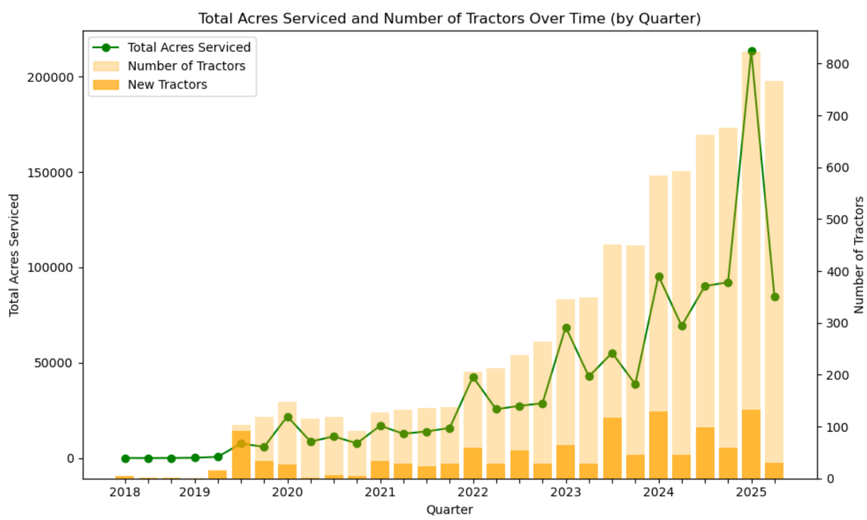
where $z_{i,j}$ denotes the elevation at pixel (i,j) and $\mathcal{N}_{i,j}$ represents the set of its eight neighboring pixels. I then aggregate this measure to the ward level by computing the mean TRI across all pixels whose centroids fall within each ward boundary.

Relative Wealth Index: I use the Relative Wealth Index (RWI) from Data for Good at Meta, which provides a model-based, geo-referenced estimate of household wealth at high spatial resolution. The index is constructed using satellite imagery and survey data and is normalized to have a mean of zero globally. I spatially join all RWI point estimates to my ward boundaries and compute the average for each ward, which I use as ward-level proxies for local socioeconomic status.

Population: I use gridded population estimates from the 2018 WorldPop "bottom-up" dataset at 100 m resolution. These data provide model-based estimates of the residential population per pixel, harmonized to UN national totals. I extract total population per ward by overlaying the ward boundaries on the raster and summing all pixel values whose area intersects each ward, yielding a ward-level measure of population size.

A.6 Additional Descriptives

Figure A3. Growth in Tractor Activity Over Time



Note:

Figure A4. Sample Comparison of Tractor Home Locations Using Census Data

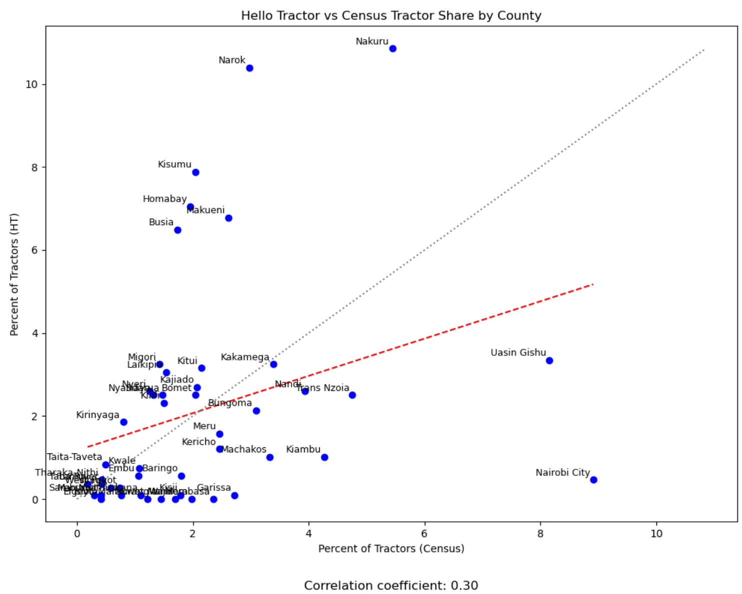


Table A1. Summary Statistics

	N	Mean	Std	Min	25%	50%	75%	Max
Jobs								
Year	917,479	2023	1	2018	2023	2024	2024	2025
Size (acres)	917,479	1.20	2.23	0.00	0.27	0.63	1.29	155.14
Av. Potential Yield (kg/ha)	864,362	4,547.45	1,160.42	-472.00	3,514.00	4,514.00	5,606.00	7,328.00
Distance to Home (km)	877,602	99.06	158.53	0.01	5.54	14.00	127.92	789.34
Tractors								
Days in Sample (in years)	1,225	1.52	1.34	0.00	0.46	1.20	2.20	5.92
Percent of Days Active	1,157	36.68	20.99	0.43	20.22	35.46	52.49	94.55
Total Acres per Year in Sample	1,225	726.40	702.49	0.18	256.14	545.26	989.10	6,593.50
Jobs per Day Active	1,225	3.90	1.61	1.00	2.82	3.80	4.76	20.67
Acres per Day Active	1,225	4.68	3.03	0.00	3.00	4.13	5.66	30.74
Assigned Home Ward Share of Nighttime Pings	1,078	66.93	25.87	11.00	44.00	69.00	93.00	100.00
Share of Jobs at Home	1,078	38.95	31.61	0.00	9.56	32.92	64.05	100.00
Av. Distance to Home (km)	1,078	68.74	95.47	0.92	8.44	23.74	89.34	554.34

Figure A5. Seasonality in Tractor Activity and Flows for 2024

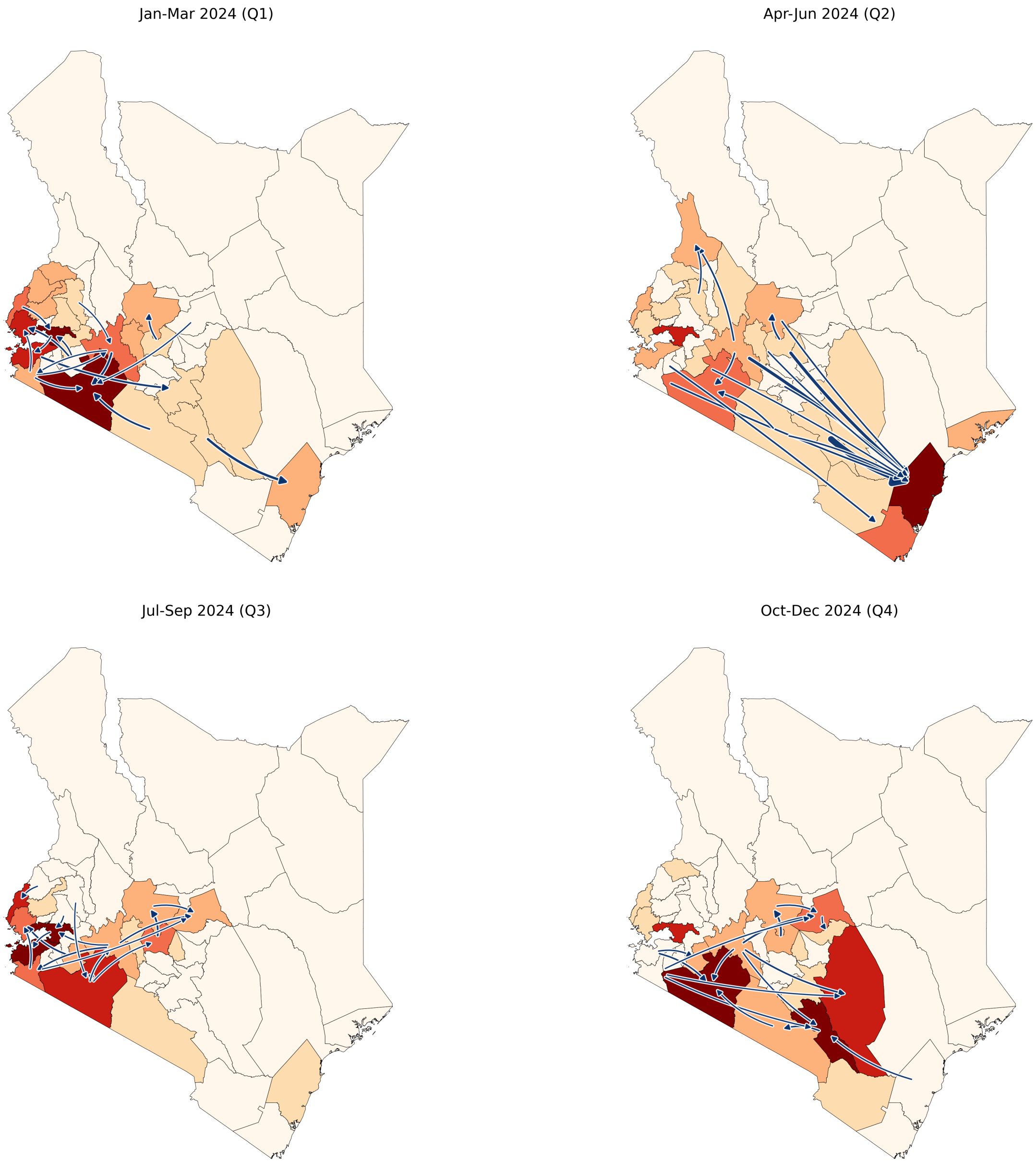
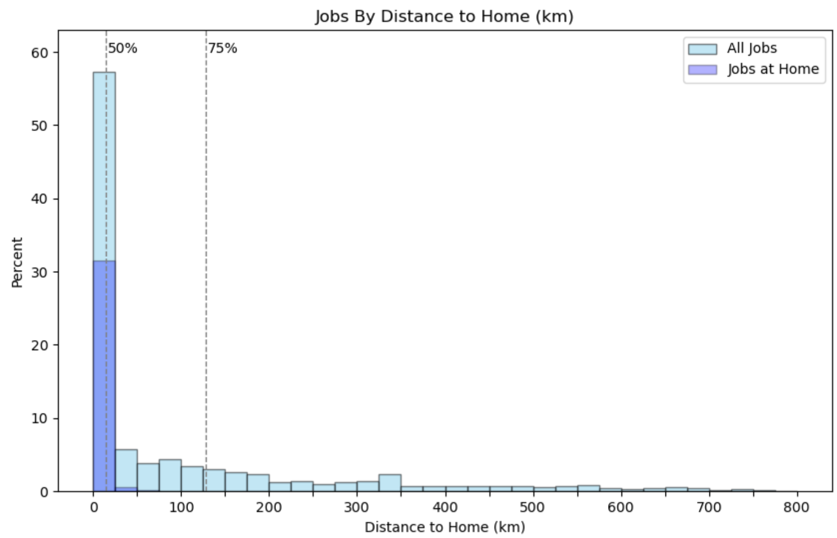


Figure A6. Gravity in Tractor Utilization



Note: Histogram of jobs by distance to home in 25km bins.

B Empirical Analysis Appendix

B.1 Data Structure Example

$$e_{it}^k = \mathbb{I}\left[\tau_{it} = k\right], \quad \tau_{it} = \lfloor (t - t_i^0)/30 \rfloor \tag{14}$$

i	t	t_i^0	$(t - t_i^0)$	τ_{it}	e_{it}^0	e_{it}^1	e_{it}^2	...
122122	2023-12-22	2023-12-22	1	0	1	0	0	...
122122	2024-01-14	2023-12-22	24	0	1	0	0	...
122122	2024-01-19	2023-12-22	29	0	1	0	0	...
122122	2024-01-29	2023-12-22	39	1	0	1	0	...
122122	2024-02-07	2023-12-22	48	1	0	1	0	...
...	...	...	...	...	...	...	...	...
122122	2024-02-19	2023-12-22	60	2	0	0	1	...

B.2 Robustness and Additional Results

Figure B1. Baseline (All Months)

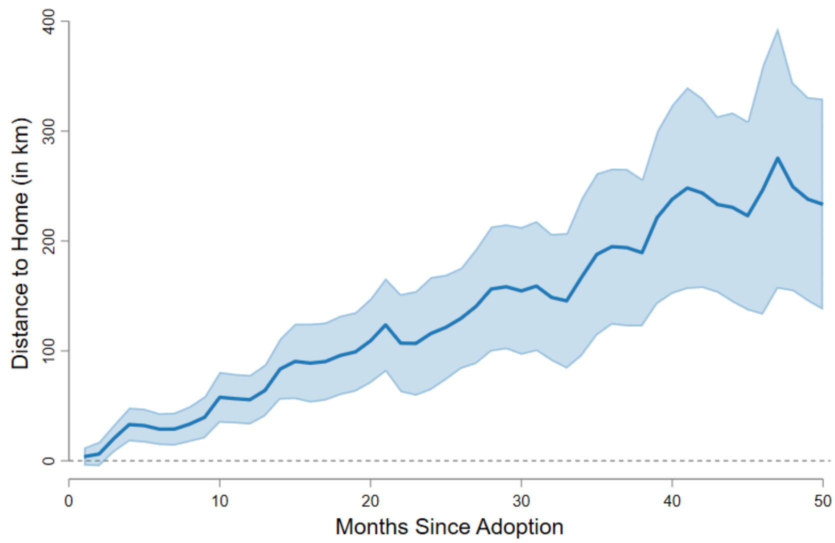


Figure B2. Potential Yields (All Months)

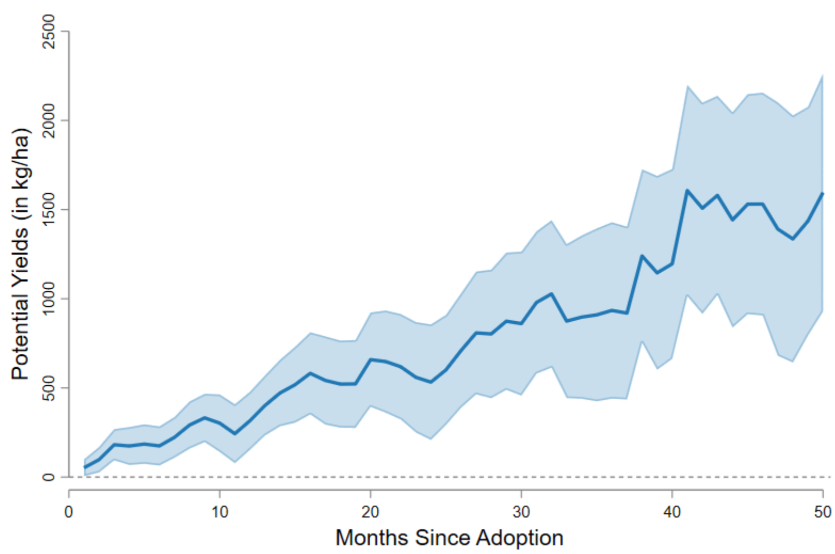


Figure B3. Tractor Mobility After Adoption by Adoption Cohort

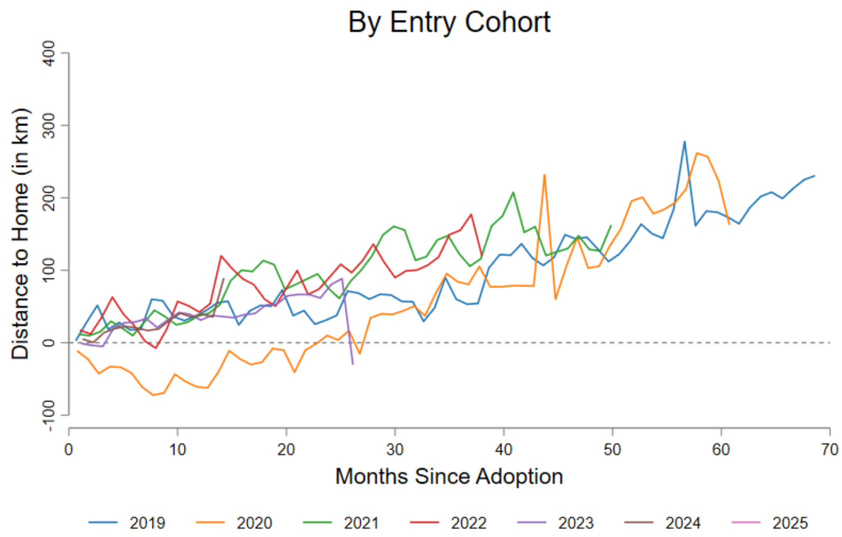


Figure B4. Tractor Mobility After Adoption Older Tractor Models

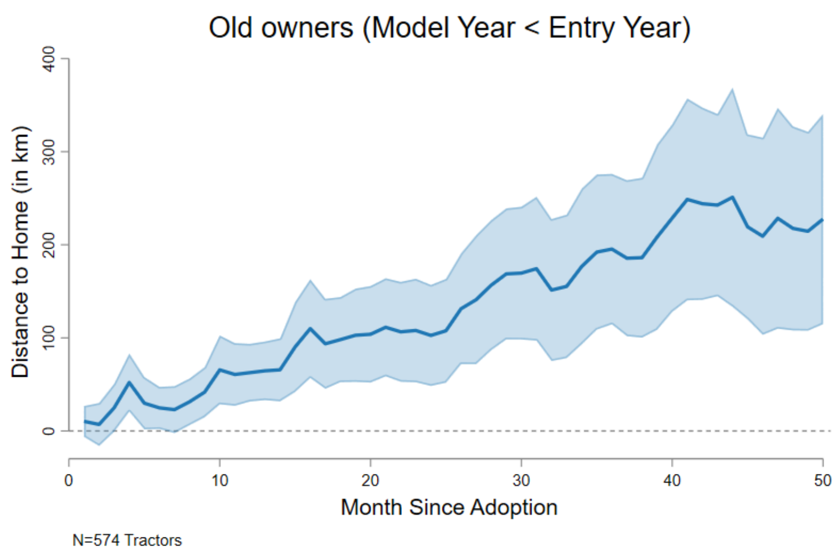


Figure B5. Robustness: High-Confidence Sample

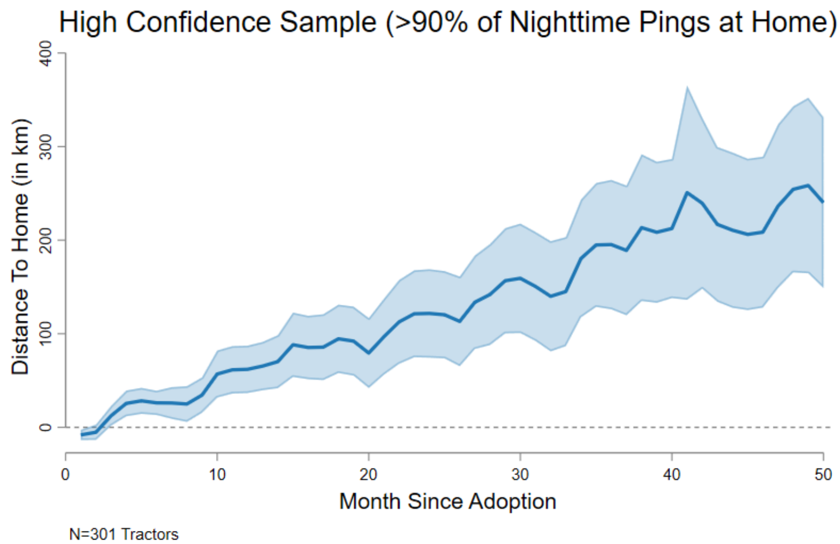


Figure B6. Propensity to Work at Home

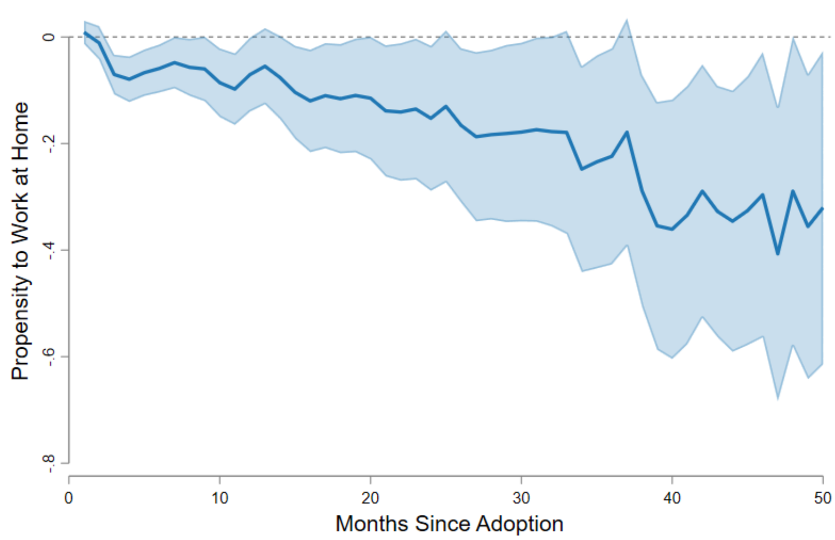


Figure B7. Propensity to Return Home After Working Outside Home Ward

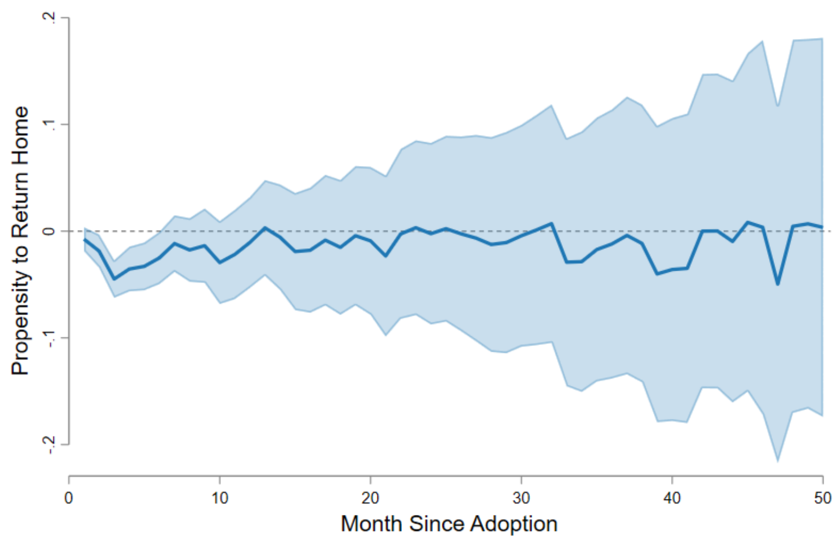


Figure B8. Probability of Being Active Over Time

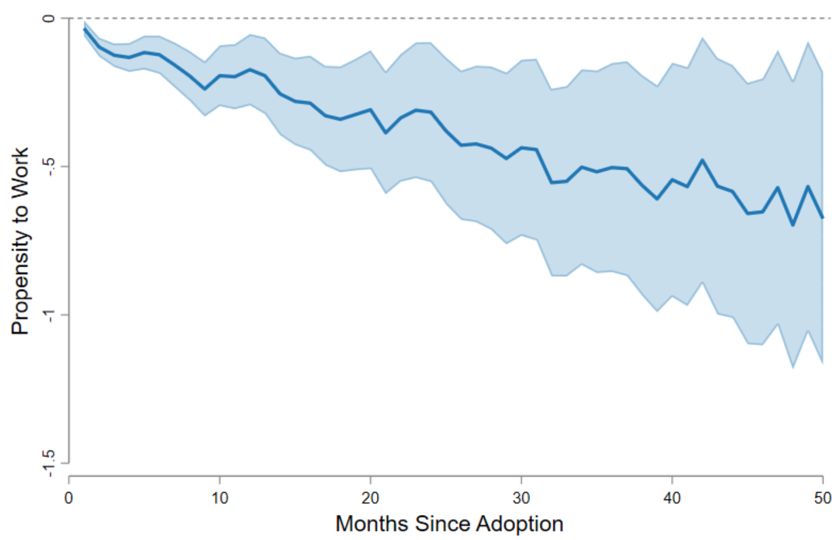


Figure B9. Number of Jobs per Day Conditional on Being Active

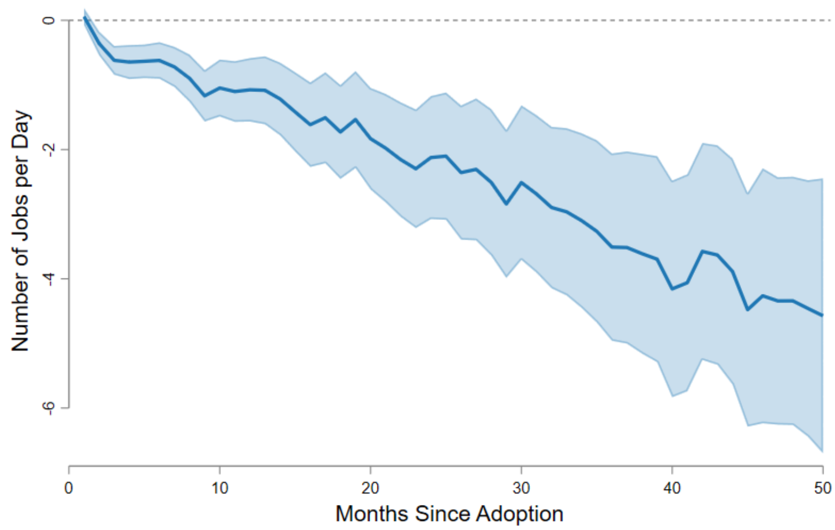


Figure B10. Effect of Tractor Visit on Start and Peak of Season NDVI