# Entry, Growth, and Exit: Theory and Evidence from Ethiopia's Cut-Flower Exporters

Ameet Morjaria* Richard Peck†

February 2019

Preliminary please do not circulate without permission‡

Industrial development is one of the most powerful poverty-alleviating forces. We study the role of cost uncertainty and learning in new industries in developing economies. With unique historical data from the cut-flower industry in Ethiopia, we document rapid entry and growth in the industry in the early 2000s. We find concurrent cost uncertainty among potential entrants and inter- and intra-firm variation in growth, consistent with learning. We go on to build a multi-agent, dynamic model that brings the production function uncertainty insights of Hausmann and Rodrik (2003) to the framework of Jovanovic (1982). Model simulations predict that the least productive firms shrink and exit each period, that firm production levels become more disbursed over time, and that firm-level growth is most volatile early in the industrial development process. We find corroborative evidence in transaction-level trade data. We conclude with an exercise that uses our model to show the potential for short-term industrial stimulus to generate long-term economic growth in the industry.

*Kellogg School of Management, Northwestern University, *a.morjaria@kellogg.northwestern.edu*

†Department of Economics, Northwestern University *richardpeck2021@u.northwestern.edu*

‡We thank Seema Jayachandran, Alessandro Pavan, Chris Udry and seminar participants at Northwestern University and EPED 2018 (Montreal) for helpful discussions. We are particularly indebted to Rocco Macchiavello with whom this project was started. Funding from the IGC and Kellogg is gratefully acknowledged. Morjaria thanks the Investment Commission, Horticultural Development Agency, Horticulture Producer Association, Ministry of Trade & Industry, EDRI and the Development Bank for hosting numerous visits. This research would have not been possible without the collaboration of many people in the Ethiopian cut-flower industry; we thank them all. For research support we would like to thank Alebel Bayrou and Tsegay Tekleselassie. The views expressed here are those of the authors alone and do not necessarily reflect those of the IGC, DFID, Kellogg or any of the organizations that funded or shared data.

# 1 Introduction

Developing countries are often characterized by large agrarian populations with substantial portion of households working in subsistence agriculture. Productivity in the agriculture sector remains low in these countries (see Figure 1). Scholars and practitioners have tried to understand the process by which vast populations can be pulled out of agriculture and into high value addition jobs. East Asian economies have successfully made this transition, and many regions of the developing world have had more limited success. Africa on the other hand has seen little growth in non-traditional agriculture, services and manufacturing (see Figure 2). As a result it is not surprising that headcount poverty numbers have remained stubbornly high in Africa. However there have been some countries that have seen rapid reduction in poverty rates since the 2000s, examples include countries like Ethiopia, Rwanda and Uganda. Ethiopia is an interesting case-in-point as its poverty numbers have come down dramatically since the mid-2000s at the same time as the government embarked on an industrial policy package to assist the structural transformation out of traditional agriculture, with the goal of creating new jobs in new industries (see Figure 3).

Job growth in medium-size and larger firms has been an important driver of poverty reduction in the 21st century. In many contexts, that industrial expansion has been aided by active industrial policy. Were there not natural barriers to the development of new industries, there would be little role for activist industrial policies. These development barriers include coordination externalities, information imperfections, or inadequate credit or risk-mitigation. Understanding and measuring these kinds of barriers is important for understanding where industrial policies work, and where they do not.

One particular imperfection–cost uncertainty that gives rise to demonstration externalities–is especially relevant but difficult to measure. In this context, firms are uncertain of the costs of production, and learn these costs as they produce. The externality comes into play when firms are able to observe the experiences of other firms in order to inform beliefs about their own production function. In this manner, early entrants into the industry generate useful

Figure 1: Agriculture shares of GDP and Employment, by region

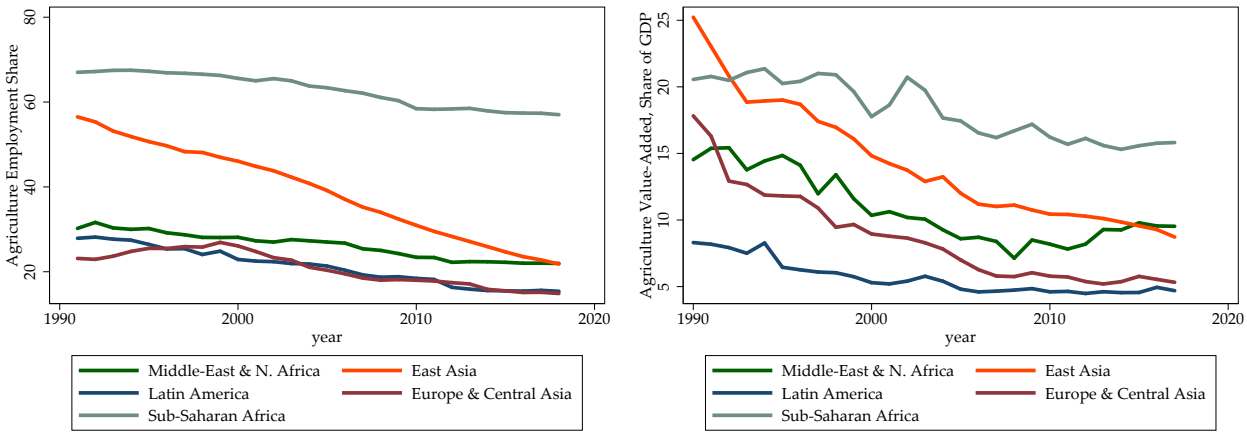
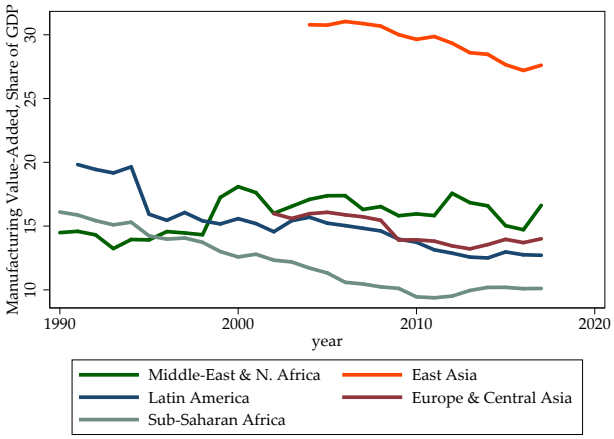



Figure 2: Manufacturing shares of GDP, by region

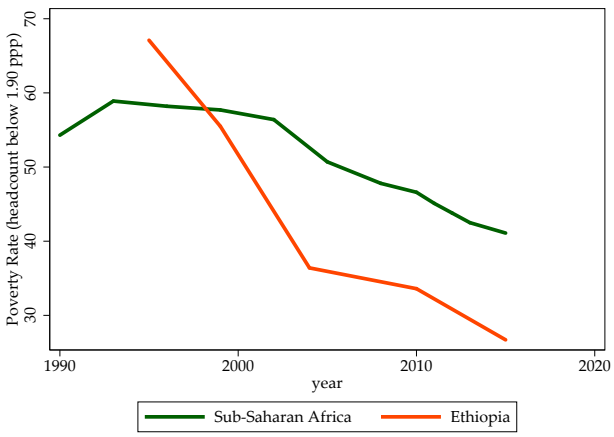



information for later participants. Such an externality offers governments justification for stimulative industrial policies.

This paper examines learning, industrial development, and policy in the context of a new industry, the cut flower exporting sector in Ethiopia. The industry grew rapidly in the mid-2000s at the same time that Ethiopia targeted credit, tax and land incentives to stimulate the industry. We examine the patterns of industrial growth and find firm behavior consistent with learning and demonstration externalities.

We push the discussion forward on three dimensions. First, we provide some stylized

Figure 3: Poverty in Ethiopia Relative to SSA




facts about the industry and introduce a dataset of firm permit applications and post-entry firm surveys which yield insights about the traits of firms that successfully enter the market. Second, we incorporate a demonstration externality (as in Hausmann and Rodrik (2003)) into the framework of Jovanovic (1982), and simulate a model of the cut flower industry that yields insights about the patterns pertinent to Bayesian learning and demonstration externalities. Finally, we examine transaction-level trade data for Ethiopian flower exporters and document patterns consistent with the learning in the simulations.

Section 2 describes the literature that we are building on. Section 3 describes our data and the cut flower industry in Ethiopia. Section 4 introduces a model of industrial development. Section 5 simulates the model in order to explore the patterns generated by a number of competing mechanisms, and compare them to the data. Section 6 simulates an industrial policy intervention, and section 7 concludes.

# 2 Literature

We are pushing forward two sets of literatures. First, we are interested in the literature begun by Jovanovic (1982) which features firms gradually learning about productivity and using

that knowledge to make entry and exit decisions. Second, we incorporate the framework of Hausmann and Rodrik (2003) to consider a demonstration externality. In this section, we briefly describe two strands of literature and empirical approaches taken by recent work, building on these seminal papers.

**Jovanovic (1982):** This seminal paper models a sector comprised of atomistic firms that are uncertain about their firm specific productivity, which is drawn from a common distribution. Firms learn their costs each period and use that information to update beliefs about productivity. This is one of the most well-cited papers in all of economics, some of the most relevant and recent work in this literature is summarized below.

Arkolakis et al. (2018) uses the Jovanovic framework in a monopolistically competitive environment. Since they have a monopolistically competitive system, they are interested in the planner problem and business stealing externalities in entry and exit. They calibrate their model with moments from Colombian plant-level data. They fit some untargetted moments to validate, and then they consider some counterfactual policies.

Albornoz et al. (2012) builds a model in which exporters can choose to export to two different destinations, which have different entry costs. The exporter has some unknown ability to export in general. It is a two stage model where the firm enters one or both markets in period 1, realizes profits and infers productivity, and then enters or exits markets in period 2 and realizes profits again. This is not exactly the learning of Jovanovic (1982), but it is somewhat in that spirit. They use Argentine customs data to test model implications.

Albornoz et al. (2016) is similar to the paper above, but in continuous time with firm-market-level profitability shifters. They conduct similar exercises in which they test predictions of the model by running regressions with Argentine customs data.

**Hausmann and Rodrik (2003)**, on the other hand, is the major paper to formalize the concept the "demonstration externality". The model progresses in three stages, where a leading firm can pay an entry cost in period one to learn the production function, and

become a monopolist in period two. Copycats enter in the final period and there are zero profits. Empirical work in the spirit of this paper has been limited. A few exceptions are below:

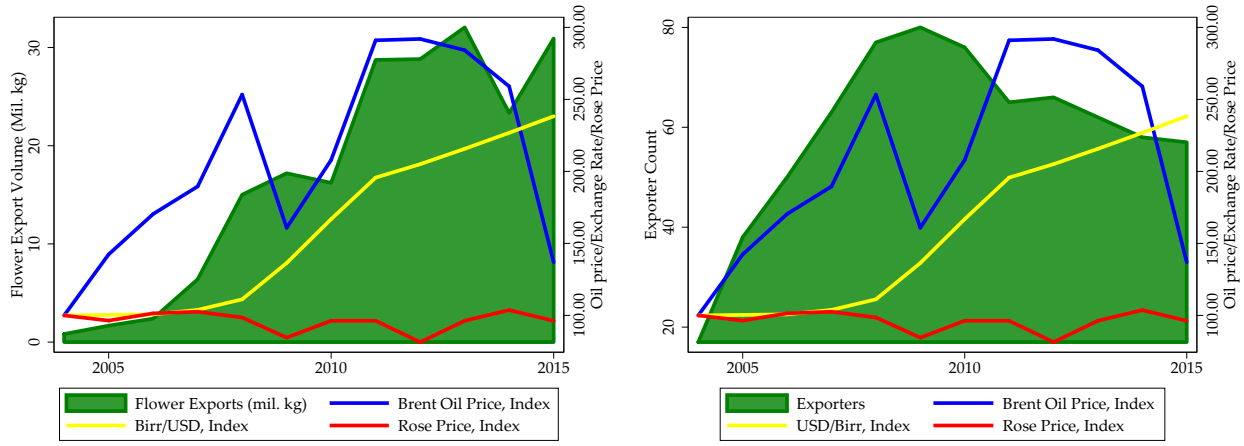
Fernandes and Tang (2014): In this model, firms are uncertain about demand in foreign markets, and update based on a number of factors including peers' experience. The model is two periods, and firms enter more when neighbors are doing well.

Artopoulos et al. (2013): examines the emergence of four different export industries and conducts interviews with firms. They document that many late entrants learned from early entrants about the industry. The finding differs from our framework, however, in that it involves learning about new technologies and adopting them, rather than resolving uncertainty about existing technologies.

Cadot et al. (2013) use transaction level export data from Malawi, Mali, Senegal and Tanzania to find correlates of survival among new entrants. They find a substantial degree of experimentation, which they argue is consistent with Hausmann and Rodrik (2003). They also find that survival is higher if there are more peer firms in that country-product-destination.

This paper is poised to contribute to each of these literatures on a number of dimensions. The Hausman-Rodrik framework has limited scope for empirical applications because it does not feature heterogeneous firms or a smooth development path. Our richer model is better able to match moments in the data to provide suggestive evidence of the presence of the demonstration externality. The context of Ethiopia is particularly well-suited to studying uncertainty about production costs that are common to the firms operating in Ethiopia. By identifying a setting in which the model fits well, we can generate further confidence about generalizing the insights of the demonstration externality hypothesis.

Figure 4: Floriculture growth in Ethiopia




**Notes:** Aggregated transaction data from Ministry of Trade & Industry. Exchange rate and oil price data are from the International Monetary Fund's International Financial Statistics. Rose prices are published by the International Trade Commission.

# 3 The cut flower industry in Ethiopia

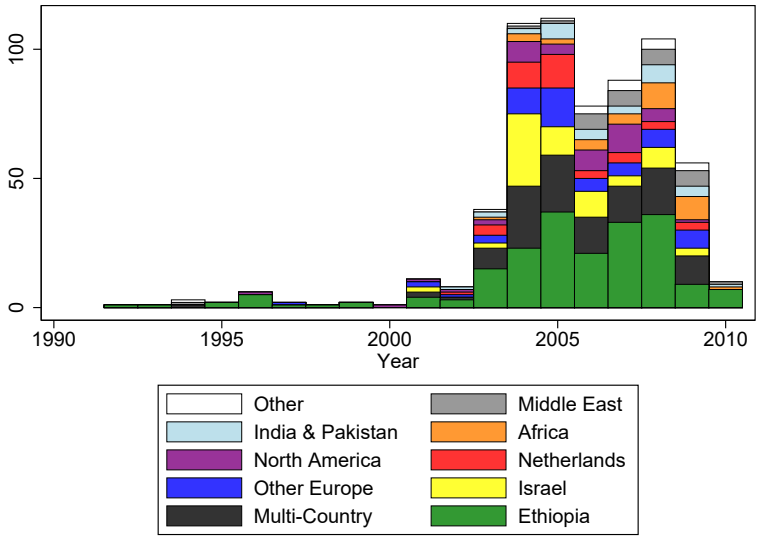
The flower export industry in Ethiopia experienced remarkable growth in the early 21st century. From 2004 to 2010, export volumes increased ten-fold.[1] Figure 4 displays this trend and the impressive entry of new firms into the market. Ethiopia feature a climate well-suited to flower cultivation. Ethiopia's southern neighbor–Kenya–is one of the top five flower exporters in the world, experienced similar rapid growth in the 1990s. This growth was aided by an Ethiopian government that offered profit tax exemptions, access to finance, and land leases which were targeted toward the industry starting in 2004. This section gives some industry background by putting our novel datasets to work.

## 3.1 Data

This paper makes use of three sets of data. The universe of investment application forms for flower farm production, a two period panel of surveys of flower farms, and transaction-level

[1]Cut flowers are generally produced for European and Middle Eastern markets, so exports are a representative indicator of industry output as a whole.

Figure 5: Flower Farm Investment Applications




**Notes:** Author's calculations on investment permit application data.

trade data that records exports for each flower farm producer.

### 3.1.1 Applications Data

First, we use the universe of Ethiopian Investment Agency investment permit applications up to 2010. As a first step toward investment in the Ethiopian floriculture sector, investors must apply to the government for an investment license. While most applicants (88%) are approved for investment, only about 20% of the 414 applications lead to operational flower farms. These applications offer an exciting insight into the cut flower industry, that few other datasets can investigate. With data on a broader pool of *potential* applicants, we can observe more about the selection process–which firms make it to market and which firms choose not to enter. Figure 5 shows how these applications increased over time, and became increasingly foreign.

Table 1 describes summary statistics for these applications. We can see about half of the applications are from solely foreign investors, and another 20% are joint ventures between foreign and domestic partners. These investments are generally new projects, but 5% of the

applications are for expansions or extensions of an existing firm. Most of the firms are in Oromia, close to the airport in Addis Ababa. About 20% of the firms lead to operational plants, and an additional 10% of the projects are in implementation phase–the project has begun, but not operational and hence exporting has not taken place. These firms are required to report their expected level of investment, expected hiring levels, and expected output level. We see that these variables are often missing–a fact that we will explore further below. In particular, over half of applicants leave the expected production question blank. Finally, we observe the financial leverage ratio and see that these investments are generally about half financed by loans.

The first trend that we are interested in noting is how the composition of entrants changes over time. Table 2 regresses indicators of foreign investors and domestic investors (note that joint venture is the excluded category) on the year and an indicator for years before the start of Ethiopia's government led push in 2004. We see that later in the sample, more firms are foreign, especially after the industrial policy push.

Next, we turn focus toward the potential entrants' expectations about inputs and outputs. We explore the correlates of these missing variables in Table 3. We see that firms far from the capital are more likely to have missing values and domestic firms are more likely to have missing values. Larger investments are less likely to have missing investment expectations. We view the prevalence of missing values as indicative that many potential entrants are uncertain about their production functions.

Finally, we examine descriptive statistics for the firms that have made it to market. Figure 6 shows the share of approved projects that were implemented by 2010.

We regress indicators of status on application characteristics in Table 4. In Columns (1) and (2), we regress on an indicator of implementation on application traits, with and without time traits. We find that domestic investors are less likely to implement, joint ventures are more likely, and larger investments are more likely to be implemented. We find that later applications are less likely to be implemented, but controlling for that, the

Table 1: Summary Statistics

| | Mean | Std. Dev. | Min | Max | Obs |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Accepted Application | 0.88 | 0.32 | 0 | 1 | 414 |
| Domestic | 0.27 | 0.44 | 0 | 1 | 414 |
| Foreign | 0.50 | 0.50 | 0 | 1 | 414 |
| Joint | 0.23 | 0.42 | 0 | 1 | 414 |
| New Investment | 0.95 | 0.23 | 0 | 1 | 243 |
| Oromia | 0.73 | 0.45 | 0 | 1 | 414 |
| Year of Application | 2005.65 | 2.83 | 1992 | 2010 | 410 |
| Began Exporting | 0.19 | 0.39 | 0 | 1 | 414 |
| Operational | 0.21 | 0.41 | 0 | 1 | 414 |
| Implementation | 0.12 | 0.32 | 0 | 1 | 414 |
| Pre-implementation | 0.68 | 0.47 | 0 | 1 | 414 |
| Investment (expected, birr) | 1929822 | 26566999 | 95 | 531150000 | 414 |
| Permanent Employees (expected) | 205.95 | 657 | 1 | 10000 | 410 |
| Temporary Employees (expected) | 402.32 | 1698 | 0 | 27770 | 409 |
| Stem production (expected, 000s) | 42778.03 | 210779 | 0.1 | 2307000 | 219 |
| Loan/Investment | 0.49 | 0.32 | 0 | 0.94 | 236 |

**Notes:** Author's calculations on investment permit application data.

Table 2: Foreign vs. Domestic over time

| | Foreign = 1 | Domestic = 1 |
|---|---|---|
| | (1) | (2) |
| Year of Application | 0.010 | 0.013 |
| | [.05] | [.055] |
| Pre-2004 | -0.772 | 1.169 |
| | [.445]* | [.434]*** |
| Obs. | 410 | 410 |

**Notes:** Author's calculations on investment permit application data.

Robust standard errors in brackets. ***p<.01,**p<.05,**p<.1

Table 3: Missing Data Correlates

| | Production (1) | Land (2) | Loan Financing (3) | Investment Type (4) |
|---|---|---|---|---|
| Missing production at capacity | | 0.721 | 2.543 | 2.251 |
| | | [.193]*** | [.259]*** | [.243]*** |
| Oromia | -0.43 | -0.311 | 0.156 | 0.269 |
| | [.187]** | [.177]* | [.308] | [.249] |
| Pre-Policy Application | -0.29 | -0.247 | -0.403 | 0.816 |
| | [.327] | [.263] | [.339] | [.51] |
| Domestic Owner | 2.62 | -0.105 | 1.716 | 1.948 |
| | [.245]*** | [.216] | [.233]*** | [.285]*** |
| Joint Venture | 0.00 | -0.242 | 0.100 | 0.311 |
| | [.182] | [.21] | [.272] | [.242] |
| Log expected investment | 0.06 | 0.010 | 0.048 | -0.161 |
| | [.042] | [.045] | [.046] | [.041]*** |
| Preimplementation Status | 0.34 | 0.031 | -0.010 | 0.023 |
| | [.172]** | [.181] | [.249] | [.252] |
| Obs | 366 | 366 | 366 | 366 |

**Notes:** Author's calculation on investment permit applications. Sample is restricted to accepted applications. Robust standard errors in brackets. ***p<.01,**p<.05,**p<.1

Figure 6: Applications by implementation status

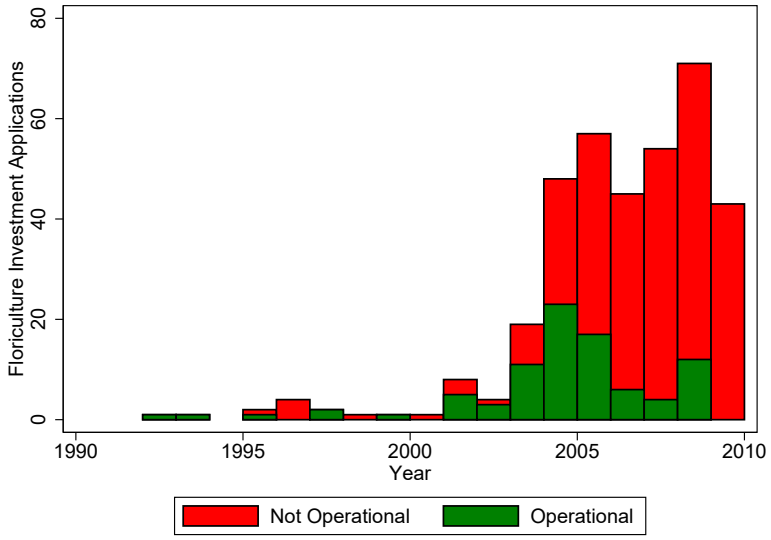


Table 4: Correlates of Implementation

| | Implemented | | Operational | | Operational | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Oromia | 0.268 | 0.716 | 0.321 | 1.006 | 0.284 | 0.982 |
| | [.314] | [.277]*** | [.392] | [.354]*** | [.51] | [.497]** |
| Pre-Policy Application | 0.622 | | 0.762 | | 0.677 | |
| | [.497] | | [.592] | | [.869] | |
| Domestic Owner | -1.033 | -0.544 | -2.205 | -0.984 | -2.488 | -1.263 |
| | [.34]*** | [.284]* | [.519]*** | [.361]*** | [.635]*** | [.522]** |
| Joint Venture | 0.251 | 0.278 | -0.047 | 0.210 | -0.880 | -0.430 |
| | [.277] | [.277] | [.302] | [.292] | [.505]* | [.468] |
| Log expected investment | 0.029 | 0.056 | 0.276 | 0.132 | -0.011 | -0.243 |
| | [.065] | [.06] | [.076]*** | [.066]** | [.156] | [.103]** |
| Year | -0.225 | | -0.325 | | -0.314 | |
| | [.07]*** | | [.076]*** | | [.142]** | |
| Sample | Approved | | Approved | | Implemented | |
| Obs | 362 | 366 | 362 | 366 | 121 | 124 |

**Notes:** Author's calculation on investment permit applications. Sample is restricted to accepted applications. Robust standard errors in brackets. ***p<.01,**p<.05,**p<.1

industrial policy is not associated with a large increase in implementation rates. A nearly identical relationship is borne out for an indicator of "operational" status. Finally, we restrict the sample to implemented projects and regress an indicator of operations beginning on these same correlates. The pattern that domestic firms are slower to operationalize is even more extreme, but the investment size ceases to be significantly correlated with operational status.

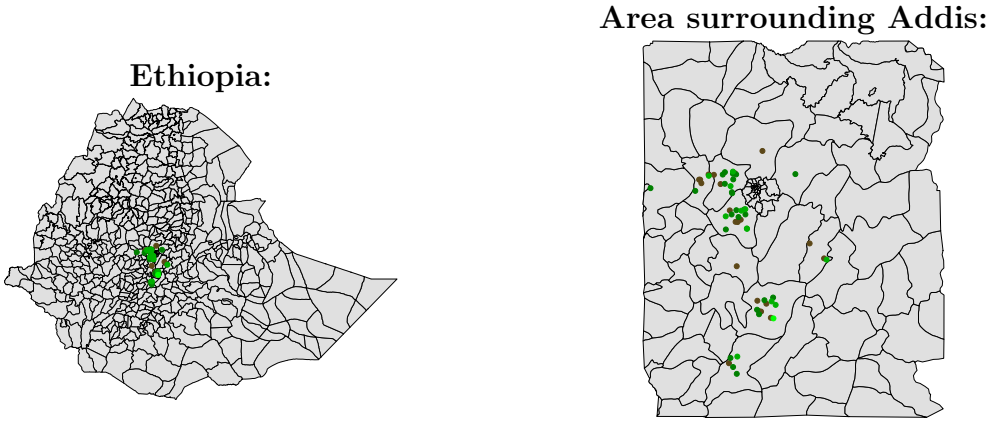
### 3.1.2 Firm Surveys

In addition to the applications data, we have access to two flower farm surveys from 2008 and 2010.[2] These two surveys allow us to have a first glance at the birth of the industry (4 years in) and a short period after (6 years in).

We can use these surveys to get additional information about the industry that will inform our modeling decisions. Figure 7 shows the placement of firms over time. We see

[2]We thank the GRIPS team for allowing us access to these two surveys.

Figure 7: Firm Placement




**Notes: Newer farms are lighter green.**

that firms are clustered around Addis Ababa, where the major international airport is. Land around the capital is abundant, and firms do not report difficulty in finding land. Figure 7 shows how this firm placement varies over time. There does not appear to be a systematic relationship between timing of entry and distance to the city, as land is abundant. The surveys offer further insight into firm placement by asking the time it takes to get products to the airport. Figure 8 reports driving distance and time to the airport. Most firms are within two hours of the airport. About three quarters of firms own their own refrigerated trucks and the remainder contract with trucking companies. Most firms ship their product through Ethiopian Airlines, a government-owned carrier that offers preferential rates to the industry. As figure 9, shows that the lion's share of costs are transport, so minimizing these costs is crucial for the firms.

Finally, the firm surveys offer some insight into the way that industrial policy induced entry. The Ethiopian government offered substantially subsidized land leases, tax holidays, preferential rates on credit, and facilitated to formation of trade groups to address the business challenges of flower exporters. In the 2008 survey, flower firms were asked why they began operating in Ethiopia. Figure 10 shows that many firms list the industrial policy as a major reason for establishing operations in Ethiopia.

Figure 8: Distance to airport

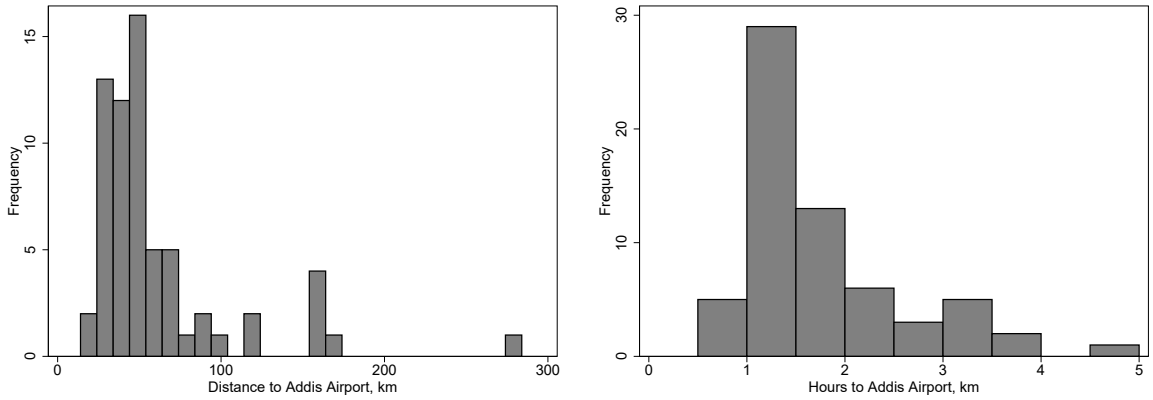



Figure 9

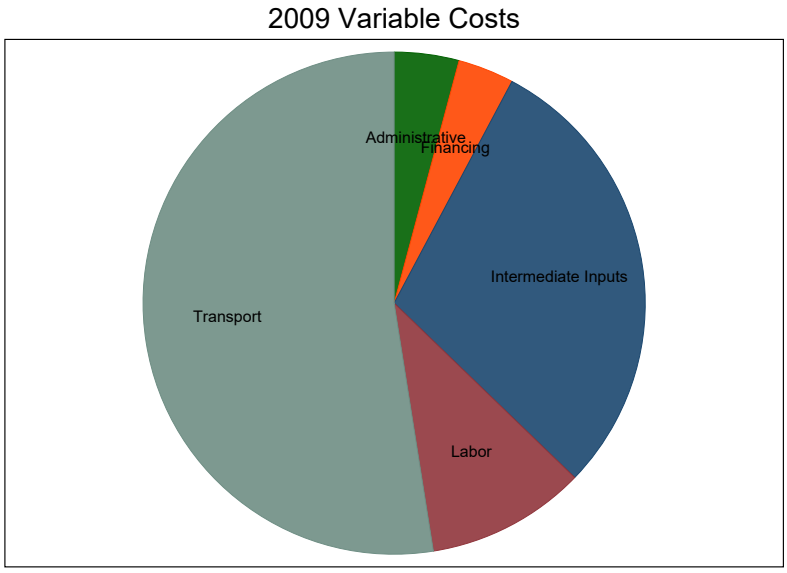
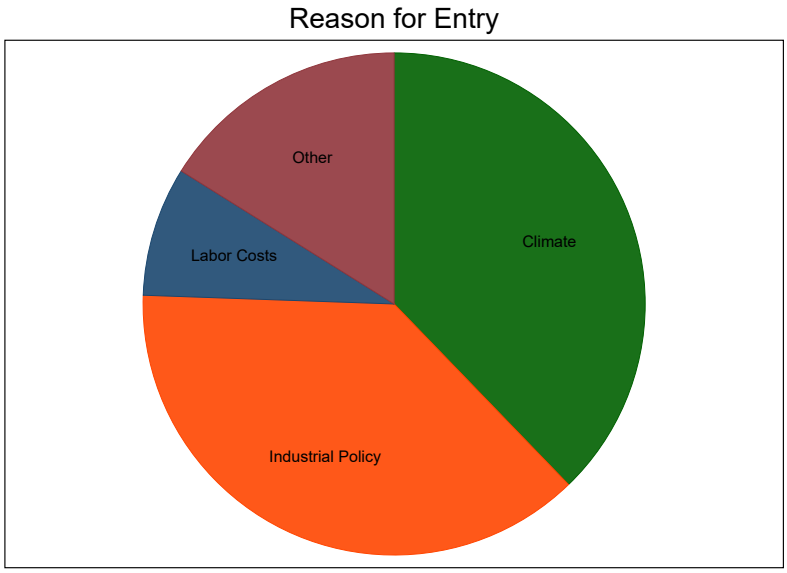


Figure 10

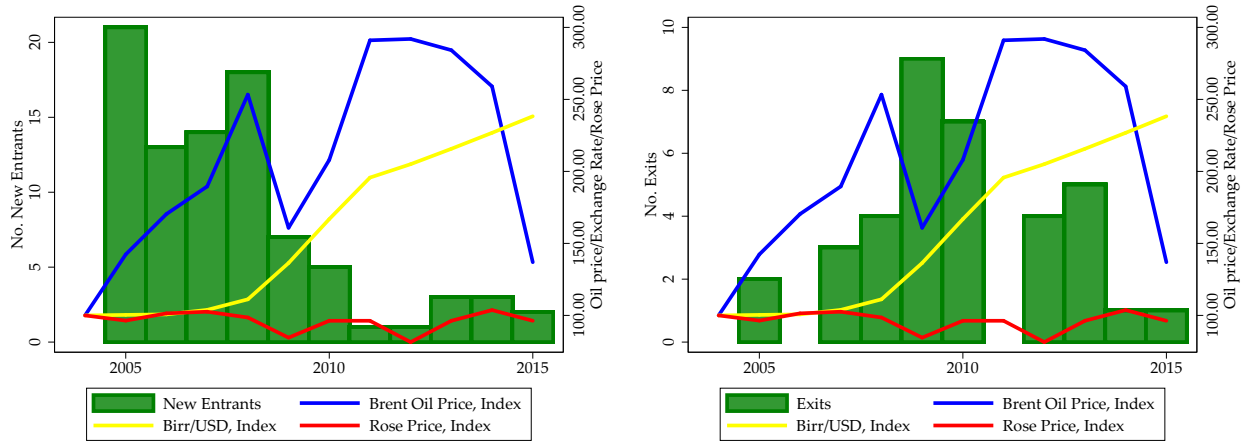



### 3.1.3 Transaction-level Trade Data

We have a transaction-level from 2004 to 2015, which reports all exports by floriculture firms. We can use these data first to assess some aggregate trends. However, there are a number of firms that exist for one quarter or for just a few transactions. At this stage we drop firms that exist for less than three quarters. In figure 11, we report entry and exit counts into the customs data. Entry is defined as the first period in which a firm exists, and exit is defined as the last period in which a firm exists. Using these definitions, we are likely to have measurement error at the beginning and end of the sample.

We see that firms are entering steadily though the period, and exits are also considerable. This is occurring as production costs are generally rising (because a large part of the production relies on imports, and transport cost also on the rise due to oil price increase), and rose prices are stable and then falling.

Figure 11: Firm Entries and Exits




**Notes:** Aggregated quarterly data from Ministry of Trade & Industry. Exit is defined as the last period that a firm appears in the dataset, and entry is the first. Exchange rate and oil price data are from the International Monetary Fund's International Financial Statistics. Rose prices are published by the International Trade Commission.

# 4 Model

In this section, we build a model that formalize our understanding of learning with demonstration externalities. We want a model that incorporates (a) firm heterogeneity and exit that we see in the data, (b) firm growth after entry, (c) uncertainty and learning about the parameters of production, (d) other shifts in price and cost over time. We will use this model to make predictions about the patterns of growth that we should see in the transaction-level trade data.

**Profits:** Each firm produces a homogenous good. At the beginning of the period $t$, each firm $i$ specifies the quantity $q_{it}$ that it will produce. Then the firm must pay cost $c(q_{it})x_{it}$ and receive revenue $p_t q_{it}$. Inverse productivity $x_{it}$ and price $p_t$ is not revealed until after production decisions are made.

The natural logarithm of prices **prices**, $p_t$ follows a random walk, so

$$\ln p_t = \ln p_{t-1} + \epsilon_t^p$$

where $\epsilon_t^p \sim \mathcal{N}(0, \sigma_p^2)$.

**Costs** are a convex, increasing function of $q_{it}$, multiplied by an inverse productivity measure $x_{it}$. Specifically, $f(q_{it})$ is of the form

$$q_{it}^{\alpha}$$

where $\alpha > 1$. Using the survey data, we estimate $\alpha = 1.35$. Inverse productivity $x_{it}$ is a function of several parameters–some of which are common to all firms, some of which are constant over time:

$$x_t = \xi\left(\theta + u_t + \zeta_i + \epsilon_{it} + C/N_t\right)$$

where

- The function $\xi(.)$ is $e^{(.)}$.
- $\theta$ is common to all firms and invariant over time. This captures the extent to which overall production in the industry is high or low cost.
- $u_t$ is a common shock to costs for all firms that is drawn from a $\mathcal{N}(0, \sigma_u^2)$ distribution and is independent over time. This shock could be something like fuel (transport) prices going up or a common weather shock.
- $\zeta_i$ is a time invariant productivity shifter that is specific to each firm. This is drawn from a distribution $\mathcal{N}(0, \sigma_\zeta^2)$. This could be managerial expertise or better supply chain networks.
- $\epsilon_{it}$ is an i.i.d. shock drawn each period for each firm from $\mathcal{N}(0, \sigma_\epsilon^2)$. It is firm specific.
- $C/N_t$ is comprised of a scalar $C$ and $N_i$, the number of firms active in the industry.

This term is meant to model a coordination externality–costs are higher when $N_t$ is low.

**Information and beliefs:** $\theta$ and $\zeta_i$ are unknown in period zero and firms gradually learn about them. Firms in the economy have common beliefs $\beta_t^\theta = (\hat{\theta}, \hat{\sigma}_\theta^2)$ about $theta$, and firm specific beliefs $\beta_{it}^{\zeta_i} = (\hat{\zeta}_i, \hat{\sigma}_{\zeta_i}^2)$ about their firm specific parameter $\zeta_i$. The realizations of $u_t$ and $\epsilon_t$ are unknown, but their distribution is known. $N_t$ is common knowledge. At the end of period $t$, all firms observe $x_{it}$.

**Static optimization:** Given beliefs $\beta_t^\theta$ and $\beta_{it}^{\zeta_i}$ and price $p_t$, firms face a decision about quantity each period. Expected profits are given by

$$\pi_i(q_{it}, \beta_t^\theta, \beta_{it}^{\zeta_i}, N_t, p_t) = p_t q_{it} - \exp\left\{\hat{\theta}_t + \frac{\hat{\sigma}_\theta^2}{2} + \frac{\sigma_u^2}{2} + \zeta_i + \frac{\hat{\sigma}_{\zeta_i}^2}{2} + \frac{\sigma_\epsilon^2}{2} + C/N_t\right\} q_{it}^\alpha$$

So profits are maximized when

$$q_{it}^* = \left(\frac{p_t}{\alpha \exp\left\{\hat{\theta}_t + \frac{\hat{\sigma}_\theta^2}{2} + \frac{\sigma_u^2}{2} + \zeta_i + \frac{\hat{\sigma}_{\zeta_i}^2}{2} + \frac{\sigma_\epsilon^2}{2} + C/N_t\right\}}\right)^{\frac{1}{\alpha-1}}$$

and expected profits are

$$\pi_i^*(\beta_t^\theta, \beta_{it}^{\zeta_i}, N_t, p_t) = \left(\frac{1}{\alpha^{\frac{1}{\alpha-1}}} - \frac{1}{\alpha^{\frac{\alpha}{\alpha-1}}}\right)\left(\frac{p_t^\alpha}{\exp\left\{\hat{\theta}_t + \frac{\hat{\sigma}_\theta^2}{2} + \frac{\sigma_u^2}{2} + \zeta_i + \frac{\hat{\sigma}_{\zeta_i}^2}{2} + \frac{\sigma_\epsilon^2}{2} + C/N_t\right\}}\right)^{\frac{1}{\alpha-1}}$$

**Exit:** Each firm has a scrap value of $W$. At the beginning of each period, a firm can permanently shut down business for a one time payout of $W$. The future stream of profits is discounted at rate $r$. Hence, the firm's value function is

$$V(\beta_t^\theta, \beta_{it}^{\zeta_i}, N_t, p_t) = \max\left\{\pi_i^*(\beta_t^\theta, \beta_{it}^{\zeta_i}, N_t, p_t) + \frac{1}{1+r}\mathbb{E}_{\beta_t^\theta, \beta_{it}^{\zeta_i}} V(\beta_{t+1}^\theta, \beta_{it+1}^{\zeta_i}, N_t, p_{t+1}), W\right\}$$

Importantly, firms are myopic with respect to entry and exit decision. This is largely a simplifying assumption, so that the firm's dynamic problem does not have to consider the

beliefs and productivities of other firms. We define $\mathcal{E}_t$ as the set of firms that have exited.

**Entry:** Each period $t$ new potential entrants arrive at rate $\lambda_t$. Each potential entrant in the set of arrivals $\mathcal{A}_t$ is assigned its firm specific productivity. Entering the market costs $\kappa_t$. Given beliefs about productivity and the number of firms in the market, firms enter the market if

$$V(\beta_t^{\theta}, \beta_{it}^{\zeta_i}, N_t + 1, p_t) > \kappa_t$$

**Beliefs and learning about aggregate productivity:** Beliefs about the common parameter $\theta$ are normally distributed and common to all firms. In period 0, beliefs are

$$\beta_0^{\theta} = \mathcal{N}(\hat{\theta}_0, \hat{\sigma}_0^2)$$

At the end of each period, firms' costs and quantities of production are common knowledge. Therefore,

$$\xi^{-1}(x_{it})$$

is observable to all firms. Hence,

$$\frac{\sum_{i \in \mathcal{I}_{\sqcup}} \xi^{-1}(x_{it})}{N_t}$$

is a signal that is informative about $\theta$. On average, the firm specific terms would cancel out if firm specific productivities ($\zeta_i$ and $\epsilon_{it}$) had a mean of zero. While $\zeta_i$ is mean zero for *potential* entrants, selective entry and exit means that that it has a non-zero mean. This non-zero mean is

$$B_t = -\frac{\sum_{i \in \left( \left[ \cup_{\tau \leq t} \mathcal{A}_\tau \right] / \mathcal{N}_t \right)} \zeta_i}{N_t}$$

We assume that this bias term is known to all agents. Moreover, the economy-wide average cost include the coordination externality cost shifter, which is known to all agents. Hence, the signal

$$\bar{\theta}_t = \frac{\sum_{i \in \mathcal{I}_{\sqcup}} \xi^{-1}(x_{it})}{N_t} - B_t - \frac{C}{N_t}$$

can be used to update beliefs about $\theta$. The firms learn according to a Kalman filter, with Kalman gain:

$$K_t^\theta = \frac{\sigma^2_{\theta,t|t-1}}{\sigma_u^2 + \sigma_\epsilon^2/N_t + \sigma^2_{\theta,t|t-1}}$$

and beliefs are updated according to the rules

$$\sigma^2_{\theta,t+1|t} = \sigma^2_{\theta,t|t-1} - K_t^\theta \sigma^2_{\theta,t|t-1}$$

and

$$\hat{\theta}_{t+1} = \hat{\theta}_t + K_t^\theta(\bar{\theta}_t - \hat{\theta}_t).$$

**Beliefs and learning about firm-specific productivity:** At the end of each period, each firm $i$ observes its own costs and $\xi^{-1}(x_{it})$. This term includes the coordination externality, the aggregate parameter, and the aggregate shock. If the firm adjusts for those factors then:

$$\bar{\zeta}_{it} = \xi^{-1}(x_{it}) - \bar{\theta}_t - \frac{C}{N_t}$$

is an informative signal about firm-specific productivity $\zeta_{it}$. Then the firm-specific Kalman gain is

$$K_t^\zeta = \frac{\hat{\sigma}^2_{\zeta,t|t-1}}{(1 + 1/N_t)\,\sigma_\epsilon^2 + \hat{\sigma}^2_{\zeta,t|t-1}}$$

and beliefs are updated according to

$$\sigma^2_{\zeta,t+1|t} = \sigma^2_{\zeta,t|t-1} - K_t^\zeta \sigma^2_{\zeta,t|t-1}$$

and

$$\hat{\zeta}_{it+1} = \hat{\zeta}_{it} + K_t^\zeta(\bar{\zeta}_{it} - \hat{\zeta}_{it}).$$

Note that we have ignored the correlation between the beliefs that is induced by the correlation of the signals. Since $\epsilon_{it}$ plays a role in the updating of beliefs about both the aggregate and idiosyncratic costs, as time goes on the distributions $\beta^\theta$ and $\beta^{\zeta_i}$ become more correlated. However, it is also the case that the variances of these variables are dropping at the same time. We ignore this correlation to simplify calculations and do not believe it would have a

significant effect on our results.

Finally, we must specify the process by which firms form their beliefs about firm specific costs upon arrival. Firms receive a signal $\phi_i$ equal to

$$\phi_i = \frac{\zeta_i}{\gamma} + \frac{1-\gamma}{\gamma} z_i$$

where $z_i \sim \mathcal{N}(0, \sigma_\zeta^2)$ is random noise. $\gamma$ parameterizes the strength for the signal. When $\gamma \to 0$, firms have no knowledge of firm specific productivity. When $\gamma = 1$, firms are certain of their relative productivity. Therefore, if a firm arrives in period $t$, beliefs before entry are

$$\beta_{it}^{\zeta_i} = \mathcal{N}(\gamma\phi_i, (1-\gamma)\sigma_\zeta^2)$$

**Summary of parameters and calibration:** This model has a number of parameters and it may be helpful to summarize them here. Table 5 collects the parameters and reports the default values which are used throughout this note unless otherwise noted. The curvature of the cost function $\alpha$ is estimated based on cost information in the firm surveys. All other values are chosen to broadly match patterns in the data. In future work, we plan to estimate these more directly using the transaction trade data.

# 5 Patterns of competing mechanisms

Given this model, how do we assess its applicability to the real world? In this section, we take the following approach – we can use the model to make predictions about the time trend of key statistics, for example, how does industry concentration change over time. Once we get a prediction from simulations of the model, we can look directly at the transactions data and decide if we see that same pattern.

Of course, the same pattern could arise from multiple mechanisms. Therefore, we simulate a number of versions of the model where parameters are changed to, in effect, turn some

Table 5: Parameters

| | | |
|---|---|---|
| $\alpha$ | Curvature of the cost function | 1.35 |
| $\{\lambda_t\}$ | Firm arrival rate | 10 |
| $\{\kappa_t\}$ | Entry cost | 1 |
| $W$ | Exit value | .5 |
| $\beta$ | Discount value | .9 |
| $\{\|\Sigma_t\|\}$ | Number of potential entrants | 10 |
| $\{p_t\}$ | Output price | 1 |
| $\gamma$ | Prior knowledge about firm-specific productivity | 0.5 |
| $\theta$ | True Theta | -1 |
| $\sigma^2_{\theta,0}$ | Uncertainty about $\theta$ | 1 |
| $\sigma^2_\epsilon$ | Variance of period-to-period idiosyncratic shock | 1 |
| $\sigma^2_u$ | Variance of period-to-period economy-wide shock | 1 |
| $\sigma^2_p$ | Variance of output price growth rate | .05 |
| $\chi$ | Coordination externality | 1 |

mechanisms on and other mechanisms off. This exercise is, in some sense, is an effort towards "identification". We are trying to learn what patterns of the data identify common learning among firms. Specifically, are there patterns that we see in the data that could arise from multiple mechanisms? Which patterns are only the trademark of firm learning?

We consider the following six simulations:

1. *Learning about common productivity.* In this scenario–beliefs about common productivity $\hat{\theta}_0 = \theta = -1$ are correct, but uncertain $\hat{\sigma}^2_{\theta,0} = 1$. The initial signal of firm-specific costs is perfect $\gamma = 1$. Prices are constant ($p_t = 1$) as are entry costs $\kappa_t = 1$. There is no coordination externality $\chi = 0$.

2. *Learning about firm-specific productivity.* In this scenario–beliefs about common productivity $\hat{\theta}_0 = \theta = -1$ are correct and certain $\hat{\sigma}^2_{\theta,0} = 0$. The initial signal of firm-specific costs is imperfect, $\gamma = 0.5$. Prices are constant ($p_t = 1$) as are entry costs $\kappa_t = 0.5$. There is no coordination externality $\chi = 0$.

3. *Learning about common and firm-specific productivity.* This is a combination of the first two scenarios. Here, beliefs about common productivity $\hat{\theta}_0 = \theta = -1$ are correct,

but uncertain $\hat{\sigma}^2_{\theta,0} = 1$. The initial signal of firm-specific costs is imperfect–$\gamma = 0.5$. Prices are constant ($p_t = 1$) as are entry costs $\kappa_t = 1$. There is no coordination externality $\chi = 0$.

4. *Observed swings in output price.* In this scenario, there is no uncertainty about the production function–$\hat{\theta}_0 = \theta = -1$, $\hat{\sigma}^2_{\theta,0} = 0$, and $\gamma = 1$. Prices follow the pattern observed in the data from 2004 to 2012. That is, they fall, then rise, then plummet in period 5, corresponding to 2009. Entry costs are constant, $\kappa_t = 1$. There is still no coordination externality $\chi = 0$.

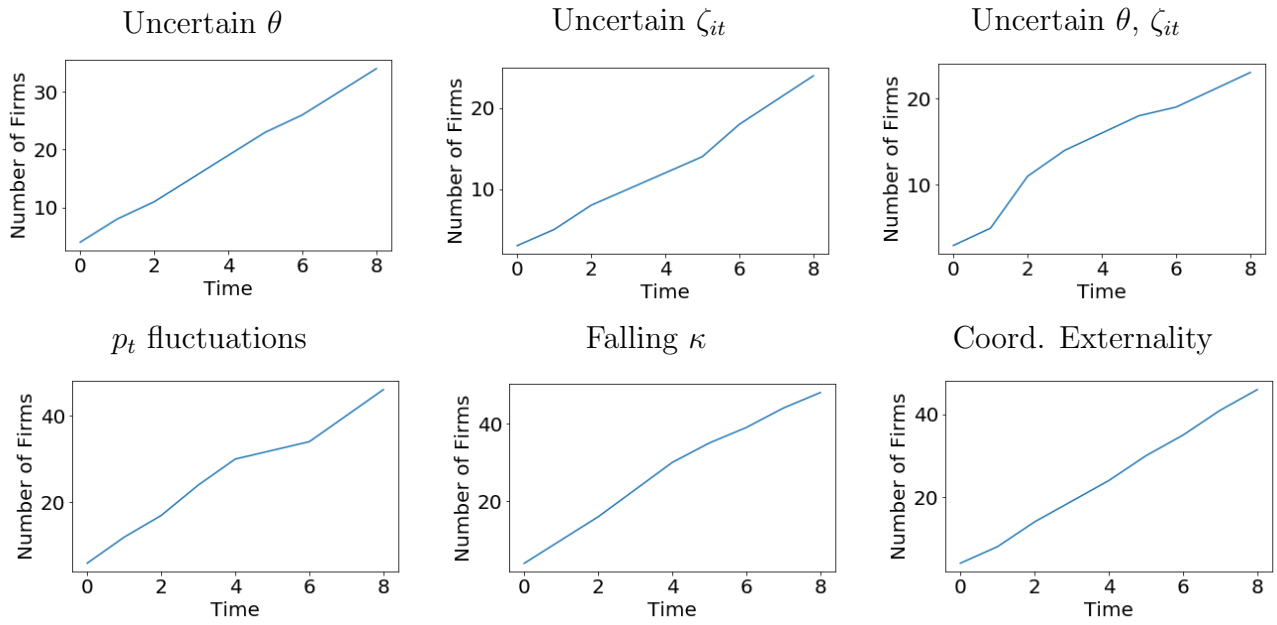
5. *A temporary fall in entry costs.* In this scenario, there is no uncertainty about the production function–$\hat{\theta}_0 = \theta = -1$, $\hat{\sigma}^2_{\theta,0} = 0$, and $\gamma = 1$. Prices are again constant. Entry costs, however, drop to 0.5 in period 1, corresponding to 2005, and remain low for 5 periods before rebounding to 1. This is meant to simulate an accommodative government policy in which the government introduces programs to make entry easier. There is still no coordination externality $\chi = 0$.

6. *Coordination Externality:* In this scenario, there is no uncertainty about the production function–$\hat{\theta}_0 = \theta = -1$, $\hat{\sigma}^2_{\theta,0} = 0$, and $\gamma = 1$. Prices are constant ($p_t = 1$) as are entry costs $\kappa_t = 0.5$. The coordination externality parameter $\chi = 1$ in this case. Thus costs are falling as more firms enter. Since we will see in Figure 12 that this leads to an increase in the number of firms over time, we can consider this simulation to shed light on any scenario in which economy wide costs are dropping over time.

Now, for each of these scenarios, we simulate the scenario 1,000 times and report average aggregate outcomes.

Our choice of patterns to focus on is largely driven by intuition about the effect of aggregate and firm-level learning on the distribution of firms. We focus on patterns that are indicative of increasing heterogeneity and decreasing uncertainty over time.

Figure 12: Number of firms over time



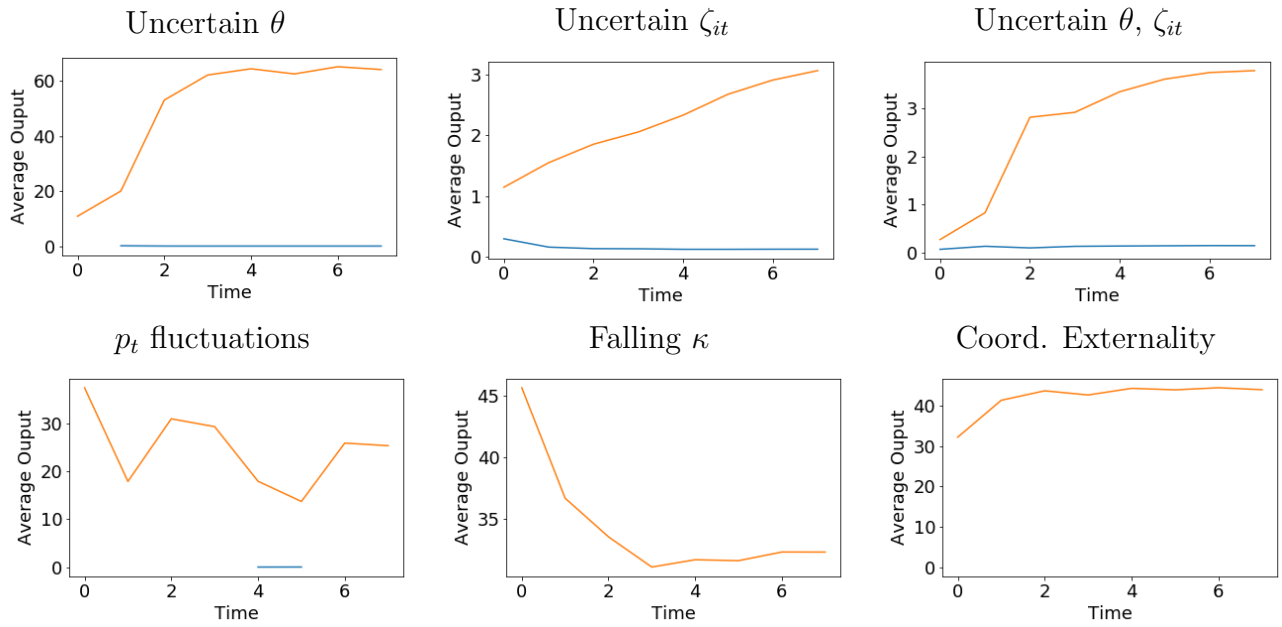

Graphs based on 10,000 simulations.

**Firm Growth:** First, however, we consider firm growth. We want to show that all of these six scenarios will deliver the main pattern that drives this paper – an explosive growth in a new sector. Figure 12 reports the average number of firms, over time, for each of these simulations. We see that in all cases, the number of firms rises. In the cases where there is uncertainty about firm-specific productivity, the number of firms is persistently lower, as uncertainty implies higher expected costs.

**Average firm size:** Next, we turn to heterogeneous firm size decisions over time and the decision to exit. We divide the firms into two groups–the firms that exit in the following period and the firms that will remain. We see in the top row of Figure 13 that the simulations with learning deliver the original Jovanovic result–remaining firms scale up production, as they receive positive signals about their productivity. The firms that are about to exit already believe their productivity to be quite low, and produce at a small scale.

The other scenario, however deliver slightly different results. The scenario in which prices

Figure 13: Average firm size



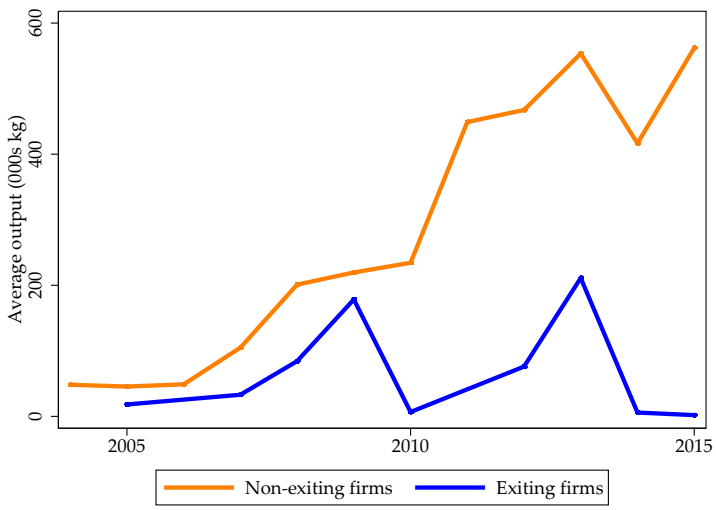

Graphs based on 10,000 simulations.

fluctuate gives the same result that only the smallest (and least productive) firms exit when prices fall. However, there is only exit when prices fall, and it cannot generate the twin phenomenon of small firms dying while remaining firms scale up does not happen. In the case of subsidized entry, there is no exit. Instead, we see that average firm size drops when entry is subsidized as lower productivity firms are able to enter. Finally, the coordination externality scenario, in which costs are falling does not deliver any exit at all, although firms do increase their scale over time as new entrants generate a coordination externality.

Figure 14 reports analogous patterns in the customs data. We see that the exiting firms are indeed the smaller firms and the remaining firms are scaling up production each period. Hence, the true data is consistent with the three learning scenarios.

**Within-Cohort Variation:** We define a cohort as a group of firms that enter at the same time. Cohort 0 enters in period zero, cohort 1 in period 1, et cetera. In a world of heterogeneous firm productivity and firm-specific learning, we expect firm scales to diverge as firms receive differing signals. Figure 15 report how the coefficient of variation grows

Figure 14: Average Firm Size–Trade Data




within a cohort, over time. As in Jovanovic, there is a significant amount of growth in the coefficient of variation when there is uncertainty about $\zeta_{it}$. Uncertainty about $\theta$ alone does not deliver this result. Nor do any of the other simulations.

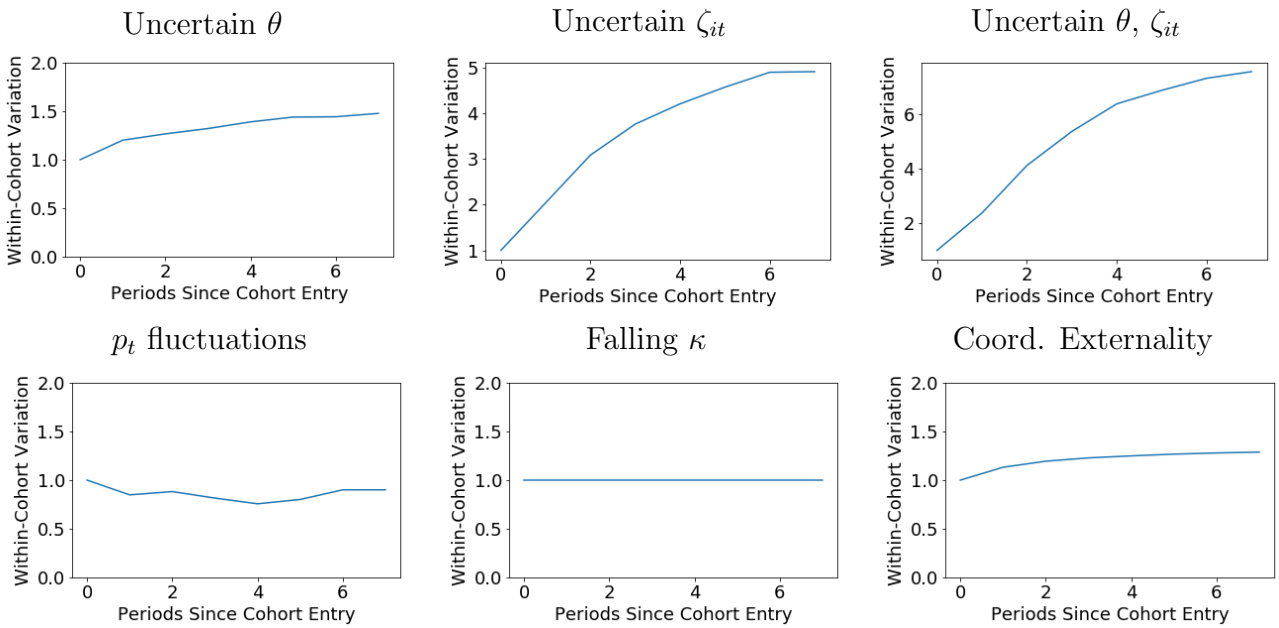
Figure 16 reports this same pattern in the data, and we see within cohort variation is rising over time. This appears consistent with the presence of firm-specific uncertainty and firm-specific learning.

**Concentration:** Of interest is the way that industry concentration changes over time. In the standard Hausmann-Rodrik model predicts that in the third period, free entry and firm symmetry dictates that the industry shifts from a monopoly to perfectly competitive. In our model, as Ethiopia is revealed to be a low cost environment in which to operate, firms continuously enter, driving down concentration. Figure 17 reports the C5 concentration ratio as well as the Herfindahl-Hirschman index in each of our simulations. We find that is is declining over time.

Figure 18 shows these measures of concentration in our data. We find that persistent entry drives industry concentration down significantly.

**Average absolute growth rate:** Finally, we calculate absolute percentage changes in

Figure 15: Within Cohort Variation




Graphs based on 10,000 simulations.

Figure 16: Within-Cohort Variation–Trade Data

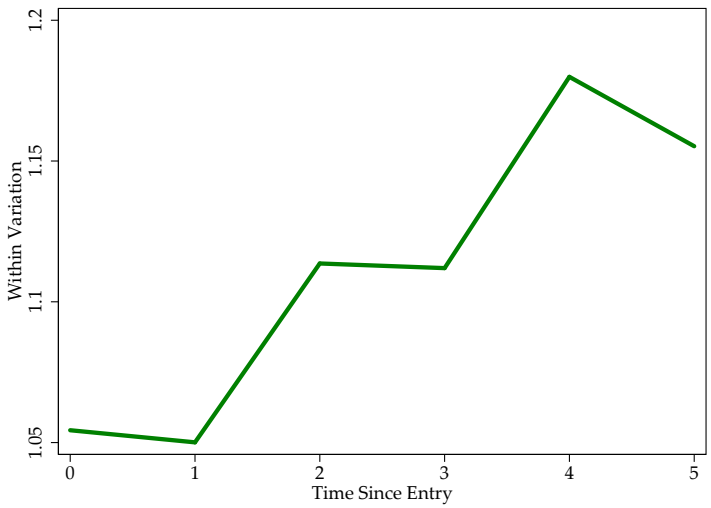
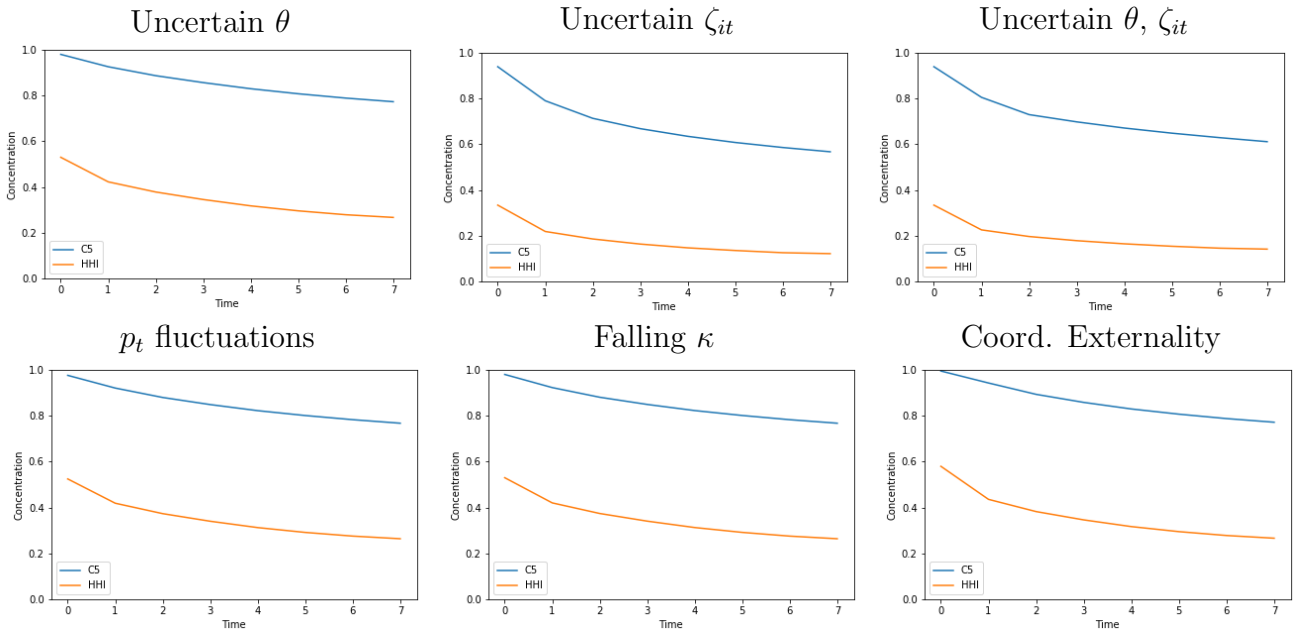


Figure 17: Concentration Measures over Time




Graphs based on 10,000 simulations.

Figure 18: Industry Concentration–Trade Data

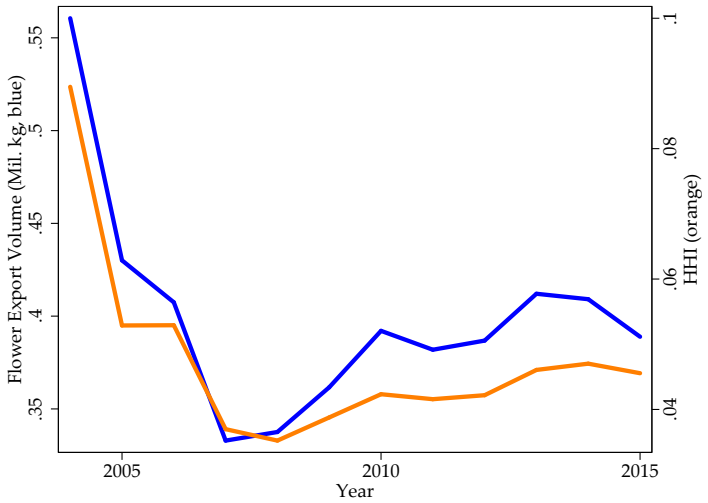
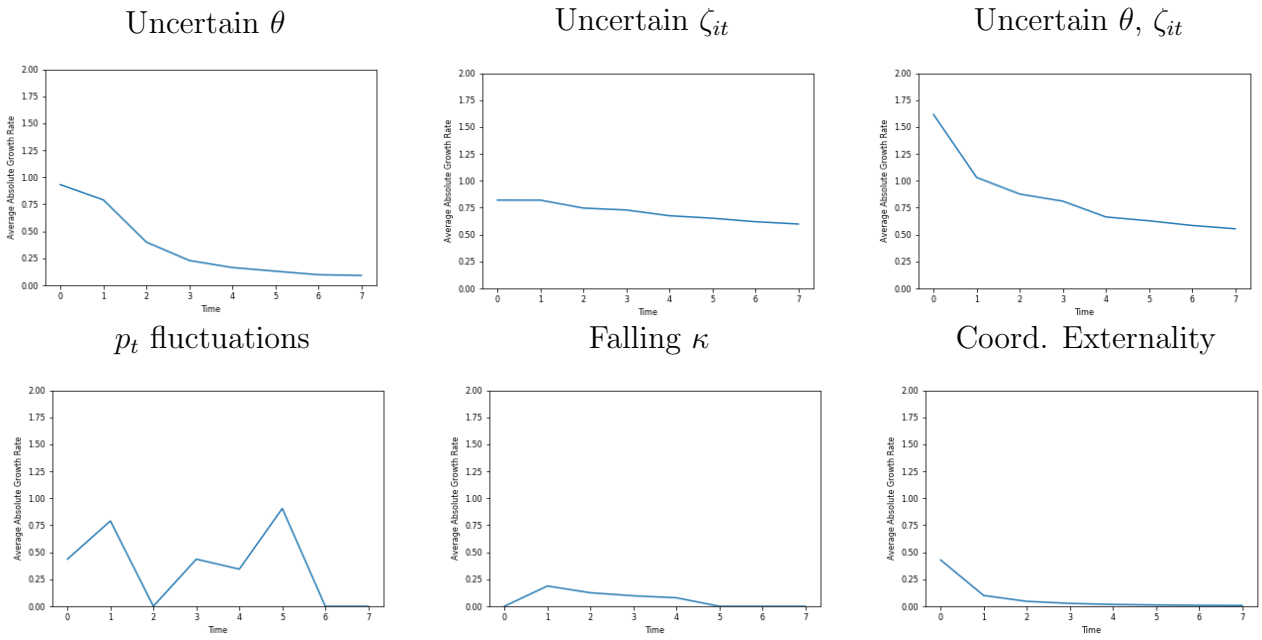


Figure 19: Absolute growth rates over time



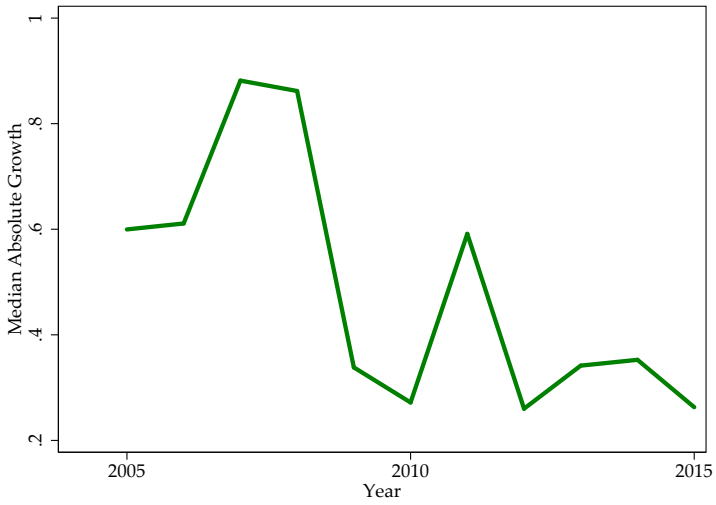

Graphs based on 10,000 simulations.

output by firm, and track how these rates of change take over time. We report this in Figure 19. We see that when there is uncertainty and learning about $\theta$, there is a lot of volatility early on, and this subsides. It subsides to zero if there is no firm-specific uncertainty, and subsides to a constant rate if there is firm-specific learning. We also get this from the coordination externality, although it is to a much smaller degree, and the growth is strictly positive in that case.

Again, we look at the same statistics for the customs data in figure 20. We see that the absolute growth rate was, indeed high early in the sample, and then it has dropped as the industry matured. This is consistent with the learning externality as well.

All in all, these results seem suggestive of the mechanisms that we posited. There is some sort of firm-specific learning going on, and common learning is likely, although it is difficult to differentiate between learning about $\theta$ and other mechanisms that have reduced costs.

Figure 20: Absolute growth rates over time–Trade Data




# 6 Policy Simulation

We want to use this model to say something about when industrial policy can be effective. Note that in the above simulations, we observed some industrial growth when there was uncertainty about productivity, even when there was no reduction in entry costs. However, these averages over 10,000 simulations masked some wildly different outcomes.

In this section, we define "industrial breakout" as the industry growing above 30 firms in 10 periods. We can then think about stimulative industrial policy as changing the probability of an "industrial breakout".

Consider the case in which $\theta = -1$, $\hat{\theta}_0 = 0.5$, and $\hat{\sigma}_0^\theta = 1$, so firms are systematically overestimating costs. If entry costs are constant at $\kappa_t = 1$, then we will see "industrial breakout" 10% of the time. However, if entry is subsidized for one period only ($\kappa_0$) then industrial breakout occurs in 42% of simulations, and a two period stimulus drives this rate up to 85%.

This quick calculation demonstrates the role of short term industrial intervention in generating longer term economic gains.

# 7 Concluding Remarks

This paper introduces unique sets of data about the development of the Ethiopian floriculture industry and builds a corresponding model of industrial development and learning. The data show substantial growth of the cut-flower industry, as well as significant exit. Applications data collected prior to entry demonstrate uncertainty about the scale and cost of production. We go on to build a model consistent with this data that bridges two important theories of industrial development –Jovanovic (1982) and Hausmann and Rodrik (2003). The model generates predictions about heterogeneity in the growth trajectories of firms, increasing dispersion of firm behavior, decreasing concentration, and decreasing volatility of firm scale decisions. We see each of these patterns born out in the data. Furthermore, the model illustrates a mechanism by which short-term industrial stimulus can have long-lasting implications for industrial development.

This is the first step in a broader research agenda to explore common-cost learning in the development of new industries. There are a number of important next steps. Using the data to better calibrate the model is chief among these. The model implies a mapping from observed export quantities to expected costs. Moreover, the model has implications about which components of expected costs should co-vary over firms within and across cohorts, and that could be used to learn more about the shape of prior beliefs. We also plan to employ a similar transaction-level data from Kenya as a useful placebo test – the Kenyan flower industry is more mature and should not exhibit the same micro-patterns.

# References

**Albornoz, Facundo, Hctor F. Calvo Pardo, Gregory Corcos, and Emanuel Ornelas**, "Sequential exporting," *Journal of International Economics*, 2012, *88* (1), 17 – 31.

_ , **Sebastin Fanelli, and Juan Carlos Hallak**, "Survival in export markets," *Journal of International Economics*, 2016, *102*, 262 – 281.

**Arkolakis, Costas, Theodore Papageorgiou, and Olga A. Timoshenko**, "Firm learning and growth," *Review of Economic Dynamics*, 2018, *27*, 146 – 168.

**Artopoulos, Alejandro, Daniel Friel, and Juan Hallak**, "Export emergence of differentiated goods from developing countries: Export pioneers and business practices in Argentina," *Journal of Development Economics*, 2013, *105* (C), 19–35.

**Cadot, Olivier, Leonardo Iacovone, Martha Denisse Pierola, and Ferdinand Rauch**, "Success and failure of African exporters," *Journal of Development Economics*, 2013, *101* (C), 284–296.

**Fernandes, Ana P. and Heiwai Tang**, "Learning to export from neighbors," *Journal of International Economics*, 2014, *94* (1), 67 – 84.

**Hausmann, Ricardo and Dani Rodrik**, "Economic development as self-discovery," *Journal of Development Economics*, 2003, *72* (2), 603–633.

**Jovanovic, Boyan**, "Selection and the Evolution of Industry," *Econometrica*, 1982, *50* (3), 649–670.