Weather-driven adaptation in perennial crop systems: An integrated study of Brazilian coffee yields *

(Job Market Paper Draft)

James A. Rising†

October 15, 2016

Abstract

Perennial crops are a third of global agricultural exports, and form an important semi-durable capital base for many developing countries. However, the dynamics of their production remain poorly understood, with the central challenge of this literature being the unobserved time-series of yields, heterogeneous ages, and farmer decisions. In this paper, I show that these issues are tightly interrelated, and require a merging between the perennial supply and statistical yield literatures. I study coffee in Brazil, exploiting variation in weather and prices as they drive changes throughout the perennial agriculture system. Extreme temperatures cause decreases in yields, but also reduce reported harvest area as farmers compensate at the intensive margin by leaving under-performing parts of their fields unharvested. At the extensive margin, extreme temperatures cause plant die-off and reduce investments in new planted area among credit-constrained farmers. As the central methodological innovation of the paper, I incorporate these effects into an integrated model, allowing the hidden effects of changing harvests and plant die-off to be distinguished. I find that the true effect of extreme temperatures on yields is 50% greater than estimated without accounting for changes in harvested area. Accounting for dynamics also reveals that the hidden costs of die-off account for half of weather-induced yield shocks, and reduce the total value of coffee assets by half. Combined, these techniques allow for a full accounting of the effects of prices, weather, and their interactions for a wide range of perennial crops.

JEL Classification: C32, O13, Q10
Keywords: perennial agriculture, environmental shocks, decision-making

*I would like to thank Solomon Hsiang, David Anthoff, Jeffrey Sachs, Wolfram Schlenker, Geoffrey Heal, Upmanu Lall, Walter Baethgen, and the participants of the Environment and Resource Economics Seminar and Climate Change Economics Lunch for invaluable comments and guidance.

†Correspondence: Energy and Resources Group, University of California, Berkeley. Email: jrising@berkeley.edu.

1 Introduction

Perennial crops account for 13% of global cropland, but produce 35% of the value of agricultural exports (Monfreda et al. 2008, Appendix 1). Low income economies are particularly reliant on perennial crops, with these accounting for an average 56% of the value of their agricultural exports and 10% of their total goods exports. Perennials have gained recent attention for the strengths that they offer under climate change, including their greater capacity for soil retention, drought resistance, carbon storage, runoff reduction, and surface cooling under intercropping (Jordan et al., 2007). While calls for research in perennial agriculture have increased, they remain understudied relative to annual crops (Devadoss and Luckstead, 2010; Wang and Alonzo, 2013). For example, the AgMIP gridded model intercomparison project includes 18 annual crops, but provides no comparison for perennial crops (Elliott et al., 2015). Below, I describe the leading difficulties in studying perennial production, and present a comprehensive solution applied to coffee in Brazil.

Perennial crops differ from annuals in economically significant ways beyond their longevity.[1] Perennial crops represent long-term investments in productive capital. These investments are under considerable risks, because of the long delays between planting and their first harvest, typically 18 - 36 months for coffee. The production of perennials also typically varies by age, where for coffee the full potential of a tree may take 10 years to achieve and declines after 15 years (see Appendix 2). As a result, farm management requires the carefully timed removal of plants with declining productivity, under the expectation of greater yields a few years later. Other management decisions may also have effects lasting multiple years, such as pruning, capping, and the introduction of similarly long-lived shade crops (Wintgens, 2009).

Another key feature of perennials is the typical availability of their data. While plantings, removals, and age-distributions are all of considerable importance to modeling perennials, data on these is typically unavailable (Elnagheeb and Florkowski, 1993). These challenges have been grappled with within the literature on short- and long-run elasticities of supply, often using modifications of the Nerlove model (Nerlove, 1958, 1979).[2] However, this discussion has largely

[1]As Brady and Marsh (2013) note, "From a modeling perspective perennial crop planting decisions may have more in common with housing and manufacturing than with annual row crop production"

[2]The Nerlove model is describe further in the literature review and analytical model sections. Conceptually, it relates changes in planted areas to prices.

ignored the supply dynamics driven by weather shocks, which have parallel data challenges.

For perennial crops, accurate estimates of yields are typically unobserved. Average yields can be computed for annuals by dividing observed production by planted area at regional scales. For perennials, planted area is rarely tracked, even by the USDA for the United States. Instead, harvest areas are recorded. Relating this data to underlying yields is an important challenge.

This paper combines insights from the perennial supply elasticity literature with the growing literature on temperature-driven yield shocks, building off of Schlenker and Roberts (2009). At their intersection, I identify important decisions and dynamics missing from both literatures. This paper sets out a seemingly simple task: to estimate a yield relationship for a perennial crop. However, the key parameters necessary for that estimation– yields, die-off, age distributions, and harvest decisions– are unobserved. Resolving these issues requires new models and robust empirical techniques.

These methods provide a new entry-point for a broad range of perennials. All perennials share "the problem of perennial" harvest yields, which I resolve. They all include intensive and extensive margin decisions, which I provide an approach to distinguish. And all of them are integrated agricultural systems, with long-term impacts of both weather and farmer decision-making.

These methods also provide a way to understand how human-environmental systems respond to shocks. Perennial crops are a form of semi-durable capital, subject to environmental risks and other factors that cause them to evolve over time. Considered broadly, planted orchards are a kind of productive capital, similar to lumber, fisheries, water resources, and air quality. When environmental shocks encroach on these systems, the response that we observe is mediated by internal responses: more selective trees are cut, regions are fished, water users are prioritized, and outdoor activities are engaged in. These systems take time to recover, as do their managers. Accounting for anything less than the combined effects of the direct shock, the short-term response to the shock, and how the long-term response interacts with additional shocks distorts our understanding of the system. A better understanding how farmers make decisions around perennial crops, and how the farmer-crop system is impacted by the environment, can inform studies of development, climate change, and food security. As such, this paper targets multiple communities; in particular, the agricultural economics and particularly the branch of its concerned with climate impacts, and the development economics communities.

The next section describes the key features of perennial dynamics and presents an analytical model of the planting and harvesting decisions, where decisions at the intensive margin are cheap but limited and at the extensive margin require years of foresight. Section three tests the implications of this model with empirical analyses, building up a fragmented picture of coffee dynamics. Section four returns to the theoretical model and directly estimates it, and presents simulation results that relate the various decisions and dynamics together. The fifth section concludes.

2 Analytical coffee dynamics

2.1 Literature Review

The original Nerlove model describes changes in planted area as a partial adjustment model, where observed areas approach "desired" planted area asymptotically (Nerlove, 1958). The desired planted area is driven by expected prices and exogenous factors. Previous studies have modified this model for perennials in a number of ways. Since average yields evolve over the lifespan of a tree, most authors modify the equations to accommodate age classes or vintages (French et al., 1985; Elnagheeb and Florkowski, 1993), although these are rarely estimated empirically (Kalaitzandonakes and Shonkwiler, 1992). If new planting replaces existing trees that are past their peak productivity, the total planted area will not expand. This insight has caused authors to distinguish planting and removals, rather than look at only aggregated changes to planted area (Hartley et al., 1987; Thang, 2011). Outside of the Nerlove literature, authors have studied the entry and exit of firms (Brady and Marsh, 2013), and provided a sounder micro-foundation to farmer decisions (Devadoss and Luckstead, 2010).

The harvesting decision plays a central role in this paper, as driven by both prices and weather shocks (Wickens and Greenfield, 1973). This connection will ultimately allow us to identify the hidden planted areas.

The elasticity literature on perennial supply dynamics falls into three broad camps, in response to the problem of managing the availability of data. Reduced-form approaches are most common, which relate contemporaneous and delayed prices to observed harvest (e.g., Wickens and Greenfield, 1973). A second option is to capture the distinction between age classes through multiple lags (Bateman, 1965). An alternative approach estimates hidden variables to represent planting and clearing, and this is found to be predictively superior to reduced form methods (Kalaitzandonakes and Shonkwiler, 1992; Elnagheeb and Florkowski, 1993). This is the approach used in this paper, combined with the delayed weather and prices used in Wickens and Greenfield (1973). Appendix 4 provides a range of supply elasticities for coffee, ranging from 0.11 to 1.0 for long-run elasticities in Brazil. These wide uncertainties demand a better understanding of the forces at work.

Within this literature, yields are typically modeled as a function of age, occasionally with random shocks (French and Matthews, 1971; Dorfman and Heien, 1989) or random evolution (Price and Wetzstein, 1999).

In contrast to these studies, an econometric yield model relates weather data, such as temperature and precipitation, with observed yields. The most advanced of these use high-resolution weather data to estimate the effect of growing degree-days (GDDs) and "extreme degree-days"(KDDs) in a non-linear fashion, and account for varying unobserved characteristics that are idiosyncratic to each region, such as management, elevation, and soil properties (Schlenker and Roberts, 2009). This approach is well-suited for annual crops when yields are observed and can be approximated as a nonlinear transformation of weather inputs. However, perennial systems including coffee plantations are much more interconnected. While yields for a given age, per unit area, are still a function of weather, that function can require multiple years of data, and be difficult to identify. Once those age-specific yields are revealed to farmers, they can adjust their effort by harvesting some areas and not others, and it is only this final form of production that is ultimately recorded. Coffee has been studied using statistical degree-day models in Colombia (Guzmán Martínez et al., 1999) and Mexico (Gay et al., 2006), however these do not address the challenge of hidden planting or the effects of weather shocks on total supply.

Weather and human decisions jointly determine changes in the planted area. Extreme temperatures, disease, and old-age reduce the harvestable area. The plantation is in constant need of new investment, to replant these lost crops and perform other forms of management. When cash is tight, these investments can be neglected, and it is exactly the loss of crops and yields that can undermine farmer solvency.

2.2 Background information

Coffee plays a vital role in many countries, providing livelihoods to 25 million inhabitants of tropical countries and supporting a $81 billion industry (Sharf, 2014), making it one of the most valuable commodities in the world. However, coffee is extremely vulnerable to climate change, perched at high elevations of hot regions. Already changes in climate are making disease outbreaks more common and shifting suitable growing regions (Guilford, 2014; Malkin, 2014).

Two species make up the vast majority of commercial coffee. *Coffea canephora* (Robusta coffee) is the hardier of the two in terms of both disease and heat resistance, and has a higher caffeine content than Arabica. *Coffea arabica* (Arabica coffee), however, is demanded for its finer taste, and remains the most widely cultivated form of coffee. Brazil produces both species.

Coffee plants require particular ranges of temperature, rainfall, and soil conditions to produce a high quality product. Arabica grows best in regions with mean annual temperatures of 18 to 22°C, while Robusta prefers temperatures between 22 and 30°C. Yields are most sensitive to the period of flowering and berry developing, with weather shocks most likely to harm the final product. High temperatures can accelerate the berry production process, but lower coffee quality (Muschler, 2001). While low temperatures are preferred, to allow the beans to accumulate flavor, frosts damage the plant, so temperatures need to remain moderate (Pendergrast, 1999). Brazil is the only large coffee-producing country prone to frosts.

Humidity should be low, but heavy precipitation (over 1400 mm per year) is important. Too much precipitation (over 3000 mm in a year), however, is harmful to the coffee plant, causes soil erosion, and encourages coffee diseases (Wrigley, 1988).

Elevation is generally considered to be a primary concern, with Arabica commonly grown above 1000 m and Robusta at lower elevations. Elevation is many countries is used to distinguish coffee quality within species as well, with quality broadly considered to increase with elevation. However, this is largely explained by the differences in temperatures: higher elevations in the tropics benefit from mountain-effect (orographic) rains but have a low enough temperature for coffee flavors to develop (Thurston et al., 2013).

Farmers incur costs throughout the life of a coffee tree. While the costs vary according to the practices engaged in by small, medium, and large farmers (see figure 1), planting costs per hectare are 50% greater than a typical year's maintenance cost, and maintenance costs can be 85% the revenue from a peak age yield (Rodriguez and Vasquez, 2009).

The capacity of farmers to make such large investments also differs by the size of the farm. While large farms have access to credit markets, nearly half of perennial farms in Brazil are less than 10 Ha. Figure 1 describes the distribution of farms in Brazil by their size, as a share of total perennial farms and of total perennial area. Total perennial crop areas have been more stable than annual crop areas, with an average 23% less than 3 Ha, 38% between 3 Ha and 20 Ha, and 39% more than 20 Ha.[3] There has been a gradual shift with medium sized farms reducing slightly in their share of the total landscape and very small (< 1 Ha) and very large ($>$ 1000 Ha) farms increasing between 1995 and 2006. This could reflect the high break-even point for revenues faced by medium-sized farms (Rodriguez and Vasquez, 2009).

When prices from Rodriguez and Vasquez (2009) are combined with the portion of farms in each size range from MSU AFRE Food Security (2015), the average prices faced by farmers is as follows. Setup costs are about 2100USD/Ha, mature plant maintenances (after age 8) is 1400USD/Ha per year, and mature plant incomes are 1700USD/Ha per year.

2.3 Decision-making on coffee plantations

Throughout the paper, the terms production, harvest, and yield will be distinguished as follows. Production is the mass of green coffee beans, measured in MT, brought to market in each year. Harvest, or harvested area, is the area of coffee trees harvested in a given year, in Ha. Yield can be computed in different ways, as discussed in section 3.4, but will always be measured in terms of $^{\mathrm{MT}}/_{\mathrm{Ha}}$. Furthermore, it is important to distinguish new planting area, when seeds or grafts are added into a plantation, from planted area, which is the entire area under cultivation irrespective of its planting year. Total planted area often is less important than bearing area, defined as the area of mature plants capable of producing yield. The opposite of new planting may take the

[3] Estimates from MSU AFRE Food Security (2015), with $^1/_3$ of the $2 - 5$ Ha group applied to areas less than 3 Ha to match the divisions made by Rodriguez and Vasquez (2009).

form of removal, the clearing of old trees to be replaced by seedlings, or of abandonment, the wholesale reduction of planted land.

The coffee farmer faces two key types of decisions: how to plant and how to harvest. The extensive margin decisions consist of new planting, removal, and abandonment of coffee acres, as studied in the coffee supply elasticity literature.

The farmer also selects which plants to harvest, in light of the observed yields, as driven by age, die-off, weather, and disease. The supply elasticity literature has assumed that this decision is driven only by plant age. In truth, it is manifestly also a matter of yield. Brazil production has had, until recently, a marked biennial cycle, characterized by "on" and "off" years. This is driven by a natural resting period for the plants, where many regions have low yields years after high-yield years. If the entire unabandonded planted area were harvested every year, irrespective of yields, we would not expect to see a similar biennial cycle amongst in harvested hectares.[4]

Furthermore, the intensive margin behaviors affects the extensive margin decision, in two ways. First, the planting, removal, and abandonment decisions are made in light of the potential for intensive-margin buffering of shocks: the true value of plants that drives these decisions must include the ability to buffer possible shocks.

The other mechanism occurs through credit constraints. Farmers try to expand their fields in periods of high prices. However, their ability to do so depends on being able to make the large investment this entails. They can finance their investments through credit markets, conditional on credit availability or through self-financing. The latter depends on the success of their most recent harvests.

Weather shocks have other long-term consequences, through plant die-off. When coffee trees die, it takes multiple years for production and yields to recover, while the plants mature.

The central driver we will consider is weather shocks. When poor weather results in low yields, it drives both an intensive margin contraction, as well as affecting the response on the extensive margin.

[4]The correlation across single year deviations of production and harvest is 0.41, significant at the 10% level with the 21 observations.

2.4 Analytical model

This section presents a simple model of heterogeneous assets, related to the Nerlove model. See Appendix 7 for a related, but more sophisticated, age-structured model, which includes optimal removals and replanting within a constant area. The purpose of this model is to describe how harvests, yields, production, and expansion are related.

Let the potential bearing area prior to year t be b_{t-1}, which includes all mature plants. Year t then produces a weather shock, w_t, which affects both plant die-off and yields.

Plant die-off is described as a fractional loss of bearing area, $d(w_t)$. At the same time, new plantings prior to year t join the bearing area. Let the age of a plant at its first yield be s, and all plants included in b_t are at least s years old. New plantings is year $t-s$, n_{t-s}, have the potential to join the bearing area in year t. However, these are also exposed to die-off in the prior s years, so that the final bearing area is year t is

$$b_t = b_{t-1}(1 - d(w_t)) + n_{t-s} \prod_{k=0}^{s-1} (1 - d(w_{t-k}))$$

New plantings are related to plant die-off and prices. Appendix 7 shows that under ecological-economically suitable conditions, replantings will also equal lost and removed plants. Removals can be described using an autoregressive relationship, so that replanting is $b_t^\alpha + \beta - b_{t-1}(1 - d(w_t))$. We also assume that new land becomes profitable under higher prices, and that this increases with existing bearing area, so the total is

$$n_t = \alpha + \beta b_t - b_{t-1}(1 - d(w_t)) + \prod_{k=0}^{\infty} \phi_k p_{t-k}$$

For simplicity, suppose that the system is static before year $t-s$, with a bearing area of b^*_{t-s}. Then

$$b_t = b^*_{t-s} \prod_{k=0}^{s} (1 - d(w_{t-k})) + n_{t-s} \prod_{k=0}^{s-1} (1 - d(w_{t-k}))$$
$$n_{t-s} = \alpha + \beta b_{t-s} - b^*_{t-s}(1 - d(w_{t-s})) + \prod_{k=s}^{\infty} \phi_k p_{t-k}$$

Combined, these produce

$$
\begin{aligned}
b_t &= b^*_{t-s} \prod_{k=0}^{s} (1 - d(w_{t-k})) + \left(\alpha + \beta b_{t-s} - b^*_{t-s}(1 - d(w_{t-s})) + \prod_{k=s}^{\infty} \phi_k p_{t-k} \right) \prod_{k=0}^{s-1} (1 - d(w_{t-k})) \\
&= (\alpha + \beta b_{t-s}) \prod_{k=0}^{s-1} (1 - d(w_{t-k})) + \prod_{k=s}^{\infty} \phi_k p_{t-k} \prod_{k=0}^{s-1} (1 - d(w_{t-k}))
\end{aligned}
$$

That is, b_t differs from b_{t-s} due to natural turn-over, profit-induced planting and abandonment, and weather-induced death. Natural turn-over captures natural death and abandonment due to changing conditions, and is represented by an autoregressive scaling term α between year $t-s$ and year t. Profit-induced planting and abandonment in year $t-s$ is assumed to follow a linear relationship parameterized by ϕ_k, which are assumed to diminish with k. Because planting recovers after s years, only die-off shocks since year $t-s+1$ affect the planting in year t.

The traditional Nerlove model describes how planted area changes with prices, but requires modifications for perennial crops (Elnagheeb and Florkowski, 1993) and weather-driven die-off. The expression used in this paper is motivated by, and then related to, traditional Nerlove dynamics.

The Nerlove relationships describe desired planted area, and an asymptotic approach to it, as driven by prices:

$$
\begin{aligned}
b_t &= b_{t-1} + a(b^*_t - b_{t-1}) \\
b^*_t &= b_0 + b_1 p^*_t \\
p^*_t &= p^*_{t-1} + d(p_{t-1} - p^*_{t-1})
\end{aligned}
$$

where b_t is the bearing area in year t, b^*_t is the asymptotically desired planted area at the price p^*_t, and p^*_t gradually updates in response to changes in the observed price p_t. The equations can be combined as

$$
b_t = \alpha + (1-a)^s b_{t-s} + \sum_{k=1}^{\infty} \phi_k p_{t-k}
$$

If the maturity age is s, so that only planting that occurs before year $t-s$ contributes to bearing area, then only the coefficients b_k for $k \geq s$ can be non-zero. Our expression is then equivalent to the Nerlove dynamics, with the added effect of die-off.

From the Nerlove equations, we can see that β also captures the rate of asymptotic approach toward the desired planted area for a given selling price. As β decreases, the rate of approach increases.

A unit area of bearing area produces yields in year t which conform to a distribution $f(y|w_t)$, where y is a random variable for yields conditioned on weather w_t. A single plant is assumed to have a yield drawn from this distribution, while larger areas on the scale of Brazilian municipalities have an empirical distribution of realized yields which conforms to $f(y|w_t)$.

Let the price of harvested yield be p_t, per MT, and the cost of harvesting be c^H per hectare. Harvesting can be thought of as starting with the most productive plants, and proceeding to lower yielding plants until the marginal costs equal marginal revenue. Harvesting will occur for all plants that exceed some threshold level, $\check{y}$, where $p\check{y} = c^H$, with yield in units of MT / Ha.

The harvested area, h_t, and quantity produced, q_t, for a given municipality in year t is then,

$$h_t = b_t \int_{\check{y}}^{\infty} f(y|w_t) dy$$
$$q_t = b_t \int_{\check{y}}^{\infty} f(y|w_t) y dy$$

for a bearing area b_t.

Appendix 6 discusses further results in the context of uncertain weather. Appendix 9 provides conditions for bioeconomically suitability, appendix 8 provides estimates for the economic value of planting, and appendix 10 describes the bioeconomic equilibrium and its transient response to weather shocks. Thang (2011) provides an overview of the literature on optimal investment decisions for perennial crops.

The purpose of the theory above is to provide intuition and testable hypotheses for the next section, and to be estimable in the integrated approach in the last section. It ignores some important features of the farming decision, including a full treatment of the effects of uncertainty (Feinerman and Tsur, 2014). If there is furthermore a potential for farmers to update their prior over the distribution of weather, they could choose to increase and decrease their planting in response to weather.

3 Empirical tests

3.1 Data

The Brazilian Institute of Geography and Statistics (IBGE) provides municipality-level production for coffee in Brazil since 1990. Nearly 2,700 municipalities produce coffee, representing an average resolution of less than 40 km. Average elevations in coffee producing counties range from 5m to 1600m above sea level. Although Robusta and Arabica are not distinguished in the data, except in the most recent year, elevation is a good proxy, with Robusta grown to 800m, and Arabica grown above 600m. This dataset allows for a broad case study of the impacts of climate change at a high spatial resolution.

The production data is combined with weather data from the Climate Forecast System Reanalysis (CFSR) since 1986. CFSR is a global weather product constructed by NOAA (Saha et al., 2010). This data product combines station and satellite measurements using weather models to produce reliable weather estimates at a high spatial and temporal resolution. The spatial resolution is .32°x.32°, a grid with boxes that are about 35 km on a side at the equator (see appendix figure 3).[5]

Growing degree-days are the integral of temperature and time, between upper and lower temperature limits, and have long been used to capture plant growth. Guzmán Martínez et al. (1999) suggest that 10°C is the appropriate base temperature for calculating GDDs for coffee. We explore a large range of minimum and maximum temperatures for GDDs, seeking the limits that provide the greatest predictive capacity. See Appendix 13 for a range of possible limits. We find that a minimum temperature of 0°C and a maximum temperature of 33°C for beneficial GDDs is optimal. This implies not only that all days over 0°C are estimated as beneficial on average, but that higher temperatures up to 33°C are progressively more beneficial. All temperatures above 33°C are combined into the measure of "extreme degree days" (KDDs). Temperatures above 33°C are not immediately detrimental, but it has a progressively smaller benefit until it becomes negative. In the preferred specification, the fitted model produces yield losses for temperatures over about 35°C. Growing degree-days are calculated as in Schlenker and Roberts

[5]The high spatial resolution is important for the mountainous areas in which coffee is grown. Where the available resolution is insufficient to capture coffee farm micro-climates, our results will be biased toward zero.

(2009), using minimum and maximum daily temperature.

The season limits for computing temperatures are also important, given the perennial nature of the plants. I explore a range of periods in appendix 12, and find that December through May provides the greatest predictive potential.

Minimum temperatures are also significant, as identified elsewhere (Jaramillo et al., 2013). Winter chill is a common requirement for perennial crops, and the increase in winter temperatures with climate change can undermine fruit production (Atkinson et al., 2013).

I also include precipitation, using the total accumulated precipitation during the same period as used for temperature. Precipitation is included as a quadratic, to capture the expectation that both too little precipitation and too much precipitation are harmfully impact yields.

Finally, I consider elevation data to distinguish Arabica and Robusta growing regions. We compute municipality elevation as the area average elevation from a gridded digital elevation map (Hastings and Dunbar, 1998).

3.2 Baseline specification

I estimate a physically-based statistical model of coffee production. The model predicts yields using a nonlinear relationship with temperature and precipitation. The model is based on Schlenker and Roberts (2009). This kind of statistical relationship is based on the biological response of coffee to temperature, but considers the *ex post* consequences of weather, farmer responses, and ecosystem and pest dynamics. If farmers are providing sufficient irrigation and shade to coffee plants, the effect of high temperatures will be mitigated beyond what biological models suggest on their own.

The form of the statistical model is,

$$\log y_{it} = \alpha_i + \gamma g_{it} + \kappa k_{it} + \mu m_{it} + \pi o_{it} + \psi o_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it} \tag{1}$$

Above and in the other models below, the observation variables and their corresponding effect estimating coefficients are:

	Var.	Coeff.
Yield	y_{it}	
Growing degree-days	g_{it}	γ
Killing degree-days	k_{it}	κ
Average minimum temperature	m_{it}	μ
Total precipitation (linear term)	o_{it}	π
Total precipitation (quadratic term)	o_{it}^2	ψ

where i indexes municipalities, t the years, and $P_{3,s(i)}(t)$ is a state-specific cubic trend to capture shifting productive capacity.

We estimate the specification above using $y_{it} = {}^{q_{it}}/_{h_{it}}$. Table 2 displays the results across all municipalities, and table 3 shows the results for the 100 municipalities with the greatest production quantities. As expected, the effect of extreme temperatures is large and negative. Growing degree-days provide a much less certain measure of yields, perhaps reflecting the counter-balancing incentives in coffee to harvest yields grown in cool areas. Rising average minimum temperatures decrease yields as well, either as a consequence of their direct biological effect or their interaction with pest species. Precipitation has a concave quadratic form. The effect of killing degree-days in the top 100 producing municipalities is only 68% of the average affect across all municipalities, suggesting that marginal lands are more susceptible to extreme temperatures.

Every additional 1000 GDDs (of which there are about 3000 on average in coffee-growing municipalities in Brazil) increases yields by about 16% (95% confidence interval 6% to 28%). Every additional 100 KDDs (an average year will have only 150 KDDs) decreases yields by 76% (71% to 81%). These values are estimated using marginal changes, so the average year is the baseline from which these percent changes are applied.

Any days in which the maximum temperature exceeds 35°C have a sharply harmful effect. As a result, even small increases in temperatures under climate change can produce large decreases in yields, particularly in regions where temperatures are currently the most productive. This is consistent with other work on the nonlinear effects of high temperatures (Schlenker and Roberts, 2009).

Figure 5 shows a graphical representation of the growing degree-day production model, with

95% confidence intervals.

3.3 Heterogeneity across regions

The model presented above describes an average effect across a wide range of environmental and climatic conditions. Robusta and Arabica coffees are both grown in Brazil, but their reported production and harvest values are combined. Elevation can be a powerful proxy to distinguish the two, with Arabica grown at higher elevations than Robusta.

We consider a form of equation 1 with municipality-specific coefficients, indexed by i.

$$\log y_{it} = \alpha_i + \gamma_i g_{it} + \kappa_i k_{it} + \mu_i m_{it} + \pi_i o_{it} + \psi_i o_{it}^2 + \epsilon_{it}$$

with the coefficients further modeled as varying linearly with elevation, using interactions:

$$\gamma_i = \gamma_0 + \beta_\gamma Elevation_i + \eta_{\gamma,i}$$
$$\kappa_i = \kappa_0 + \beta_\kappa Elevation_i + \eta_{\kappa,i}$$
$$\mu_i = \mu_0 + \beta_\mu Elevation_i + \eta_{\mu,i}$$
$$\pi_i = \pi_0 + \beta_\pi Elevation_i + \eta_{\pi,i}$$
$$\psi_i = \psi_0 + \beta_\psi Elevation_i + \eta_{\psi,i}$$

The parameters that mediate this sensitivity– the positive effect of GDDs and the negative affect of KDDs– are shown in table 4 and in a graphical form in figure 6. The estimate in table 4 is reported both for continuous elevation, and for an indicator of elevations above 800m. To make the two estimates comparable, elevation is divided by 470m, the difference between the average elevation of municipalities below 800m and those above.

We find that temperatures above 33°C at 1000m above sea level are five times as damaging as they are at 250m. These results likely reflect underlying biological differences: Arabica, grown at higher elevations, is much more sensitive to temperature than Robusta.

3.4 The problem with perennials

Since perennial crops are not planted every year, their total area is rarely recorded. Instead, yields are often reported as the quantity of production, divided by the area harvested. However, this represents a distorted perspective on yields. A comprehensive measure of yield would compare production in year t to the potential harvestable area in year $t - s$, with s is the age of the first harvest. In year $t - s$, all of the planted intended to be harvested in t would exist as plants or seeds in the ground. Should any of these die in the intervening s years, these losses should be reflected in decreases in comprehensive yield.

The typically reported measure of yield conflates three sources of yield shocks: decreases in production per plant, decrease in area harvested, and loss of plants. Consider a simple model for log yield as a functions of weather, w. Let the area of mature plants be m_t. We can define,

$$\begin{aligned}
\log {}^{q_t}/_{m_t} &= \bar{y} + a(w) && \text{instantaneous yield} \\
\log {}^{m_t}/_{b_{t-s}} &= b(w) < 0 && \text{plant die-off} \\
\log {}^{h_t}/_{m_t} &= c(w) < 0 && \text{harvest selection}
\end{aligned}$$

Instantaneous yield is the aggregate productivity, per unit of potentially harvestable area. This potentially harvestable area can differ from the bearing area s years ago due to die-off in the intervening years. Finally, harvested area can be less than the the potentially harvestable area if there is harvest selection.

From these, we can see that the comprehensive yield is the instant yield plus die-off, $\log {}^{q_t}/_{b_{t-s}} = \log {}^{q_t}/_{m_t} + \log {}^{m_t}/_{b_{t-s}} = \bar{y} + a(w) + b(w)$, where both $a(w)$ and $b(w)$ are negative in an unfavorable year. We can also see that conventional perennial yield is the instant yield artificially inflated by the harvest selection, $\log {}^{q_t}/_{h_t} = \log {}^{q_t}/_{m_t} - \log {}^{h_t}/_{m_t} = \bar{y} + a(w) - c(w)$.

Throughout this section of the paper, we use single-year corrected yields. That is, if harvested area is lower in year t than year $t-1$, then the harvested area in year $t-1$ is used for computing $y = {}^{q_t}/_{\bar{h}_t}$. This decreases yields in about 25% of observations. Figure 7 displays the empirical cumulative distribution of yields by each of these methods, and according to the integrated analysis in the section 4.

Table 5 shows the same specification as above, but columns 2 and 3 estimate the effects on log production and log harvested areas. There is a large and statistically significant negative effect on harvested acres due to killing degree-days. This suggests that in hot years where the crop is damaged, the plants are simply not harvested as fully. As a result, the actual damaging effects of high temperatures on yields are likely to be greater than estimated using $y = q/h$. The yield numbers hide the fact that unproductive plots in poor years can be left un-harvested, causing both total production and harvested acres to decrease and produce a counter-balancing effect on the dependent variable.

3.5 Direct hysteresis from weather shocks

We are also interested in the long-term effect of weather shocks. To explore these, we estimate

$$\log h_{it} = \alpha_i^H + \gamma^H g_{it} + \sum_{d=0}^{3} \kappa_d^H k_{i,t-d} + \mu^H m_{it} + \pi^H o_{it} + \psi^H o_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it}^H$$

$$\log q_{it} = \alpha_i^Q + \gamma^Q g_{it} + \sum_{d=0}^{3} \kappa_d^Q k_{i,t-d} + \mu^Q m_{it} + \pi^Q o_{it} + \psi^Q o_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it}^Q$$

As shown in figure 8, weather shocks result in an effect in both the current year and the following year.

Fields recover two years after a weather shock. The effect of weather shocks can be decomposed into three components. The contemporaneous effect of KDDs is a loss of about 0.13 % of the harvested area per degree day, compared to the average year. The effect comes the effect of significant damage with fields left unharvested due to depressed yields. A shock a year ago results in about half of that effect, 0.06 %, in the next year's harvested area. This represents the portion of the KDD effects which represent multi-year damage.

Production shows an even more pronounced effect. The loss from a contemporaneous shock is 0.36 % per degree day. This combines yield losses, physical damage, and reduced harvesting area. The effect from a shock in the previous year has a similarly large effect, with a point estimate of 0.28 %. This suggests that a large portion of the production losses represent physical damage.

No significant impact is observed from shocks more than 2 years in the past. This is consistent with a new seedlings recovering the yields of their destroyed parents after 24-36 months, depending on how early the plant is killed and replanted in the first season. This period is somewhat shorter than expected, with the first year of yields for a newly planted typically occurring in the third year. There are a few possible explanations. First, coffee plants can provide their first crops in as little as 9 months if they are grafted rather than planted. It could also indicate that we are not detecting true death, but simply a two-year-long damage to the plant. However, we will refer to this effect as die-off, since it follows that expectation fairly closely.

3.6 The investment decision

Under a similar approach, we study the decision to invest in new land, in light of variation in weather.

We use the international price, according to the World Bank's Commodity Price dataset (The Pink Sheet). The World Bank reports a coffee price indicator for Arabica averaging ex-dock market prices from New York and Bremen/Hamburg, and for Robusta averaging ex-dock market prices from New York and La Havre/Marseilles. For most of the period of interest, these values correspond closely with prices paid to growers, according to the International Coffee Organization (correlation = 0.91 for Arabica and 0.79 for Robusta). Of interest here, they also extend before the ICO and production data, allowing multiple lags of price to inform changes in harvested area.

We estimate the effect of lagged prices on harvest, as follows:[6]

$$\log h_{it}/h_{i,t-1} = \sum_{l=1}^{4} \upsilon_l \log c_{t-l} + \alpha_i + \gamma g_{it} + \kappa k_{it} + \mu m_{it} + \pi o_{it} + \psi o_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it}$$

The largest effect from prices is from two years prior. A 1% increase in price results in a .1% increase in harvested area, two years later. There is also a negative effect associated with an

[6] This and the following regressions describe differences in log harvest. While p_{t-1} may be a function of h_{t-1}, there is no endogeneity ($\mathbb{E}\left[p_{t-1}\epsilon_t\right] = 0$), so long as there is no autocorrelation in the error term.

increase in prices in the previous year. This may a reflection of more intensive harvesting to take advantage of elevated prices, which increases the denominator in the dependent variable.

We ask next if changes in harvested area responds to interest rates. Expanding production requires considerable investments and capital. The availability of capital investment depends on the interest rate available in the whole of the Brazilian economy.

The World Bank provides a yearly real interest rate, since 1996. We prefer this to the Brazil Central Bank's SELIC rate, Bank's overnight interest rate. Interest rates fluctuated wildly during the early period of our data, prior to the introduction of the Brazilian Real, in 1994. Records of the SELIC rate extend further back in history, into the 1940s. However, the period prior to 1997, when the World Bank interest rate becomes available, featured rates that were almost exclusively above any seen in the World Bank timeline.

We add this to the specification above,

$$\log h_{it}/h_{i,t-1} = \sum_{l=1}^{4} \rho r_{t-l} + \upsilon_2 \log c_{t-2} + \alpha_i + \gamma g_{it} + \kappa k_{it} + \mu m_{it} + \pi o_{it} + \psi o_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it}$$

The results are shown in table 7.

The result is that higher interest rates depress expansion, with the greatest effect for rates two years before the year of interest. There is, again, an effect of the opposite sign in the previous year. We interpret this as a higher opportunity cost to harvesting, driving down the denominator.

Finally, we look for interactions between weather shocks, prices, and interest rates. We want to understand if a weather shock harms a farmers ability to replant, in those years when the interest rate is high. To do so, we consider the full set of interactions including delayed extreme temperatures. Delayed extreme temperatures are meant to capture losses that make self-financing of new planting more difficult. The results are shown in table 8.

$$\begin{aligned}\log h_{it}/h_{i,t-1} = &\beta^r r_{t-2} + \beta^c \log c_{t-2} + \beta^{rc} r_{t-2} \log c_{t-2} + \beta^{rk} r_{t-2} k_{i,t-2} + \beta^{ck} \log c_{t-2} k_{i,t-2} + \\ &\beta^{rck} r_{t-2} \log c_{t-2} k_{i,t-2} + \alpha_i + \gamma g_{it} + \kappa k_{it} + \mu m_{it} + \pi o_{it} + \psi o_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it}\end{aligned}$$

Here, prices only matter when combined with a weather shock in the planting year. Interest rates always constrain new planting, but do so to a greater extent if there is a shock in the planting year and the farmer cannot self-finance.

4 An integrated model

The empirical findings support our theoretical model above: weather affects yields, but it also affects harvested area. In addition, we see that beyond the scope of the original model, prices, interest rates, and yields all interact in determining new planted area.

We now estimate the combined model for coffee. We will use the expressions from section 2.4, with the age of maturity, $s = 2$.

We will estimate the model using

$$f(y|\mu(w_t), \sigma) = \text{Uniform}(y|\mu(w_t)(1-\sigma), \mu(w_t)(1+\sigma))$$

with a cumulative distribution, corresponding to the harvested portion, $\lambda(y|w_t)$.

$$\lambda(y|w_t) = F(y|\mu(w_t), \sigma) = \begin{cases} 0 & \text{if } y < \mu(w_t) - \sigma \\ \frac{y-\mu(w_t)+\sigma}{2\sigma} & \text{if } y \in [\mu(w_t) - \sigma, \mu(w_t) + \sigma] \\ 1 & \text{if } y > \mu(w_t) + \sigma \end{cases}$$

To enforce smoothness on $\lambda(w_t)$ during the estimation while avoiding values outside of [0, 1], I apply an inverse logit transformation:

$$\tilde{u}(w_t) = \frac{1}{1 + e^{-(4\lambda(\tilde{y}|w_t)-2)}}$$

The same is applied to the effective yield, $\int_{\tilde{y}}^{\infty} y f(y|w_t) dt$.

Based on the observations above, only the price 2 years before harvest is included.

4.1 Hypotheses and assumptions

We use the expression in equation 1 to describe yield, excluding an effect from time and including the effect of elevation:

$$\log y_{it} = \upsilon + \gamma g_{it} + \kappa k_{it} + \xi k_{it} e_i + \mu m_{it} + \pi o_{it} + \psi o_{it}^2$$

We hypothesize that $\gamma > 0$ (GDDs increase yields), $\kappa < 0$ (KDDs decrease yields), $\mu < 0$ (average minimums decrease yields), and $\pi > 0$ and $\psi < 0$ (precipitation effects are concave) in the expression above.

Die-Off is driven only by extreme temperatures, as

$$\log d_t = \delta k_t$$

We will assume that $\delta < 0$.

We will also assume the autoregressive factor, $\alpha \in [0, 1]$, the proportional range of yields, $\sigma \in [0, 1]$, and the effect of prices, $\omega > 0$, and the harvest cost, $c^H > 0$. The portion unharvested will be $\check{\lambda}(t) \in [0, 1]$ by construction.

We will further assume that observations of harvest and production are made with geometric errors

$$\log \hat{h}_t \sim \mathcal{N}(\log h_t, \sigma_h) \tag{2}$$

$$\log \hat{q}_t \sim \mathcal{N}(\log q_t, \sigma_q) \tag{3}$$

We can express planting as a hidden variable, expressible entirely in terms of known variables and parameters, and compute this as an intermediate variable:

$$b_t = \theta h_{t-2}^{\alpha}(1 - d(w_{t-1}))(1 - d(w_t)) + p_t^{\phi} \lambda(\check{y}|w_t)^{-\alpha}$$

4.2 Integrated modeling approach

Estimating the combined effects of weather shocks on yields, harvest decisions, die-off, and new investment requires nonlinear methods. Moreover, average yields and bearing area are latent variables, which are neither directly observed nor simply translated into observed quantities. Knapp and Konyar (1991) address these issues for a Nerlove model using a Kalman filter, which handles linear estimation of systems with latent variables. To handle the nonlinear case, we apply a related set of computational Bayesian methods.

Each region is estimated independently, using equations 3 and their supporting expressions. We perform the estimation using a Markov Chain Monte Carlo (MCMC) approach to compute draws from the posterior distributions of each parameter. Each region includes 22 years of data, resulting in 20 observations. There are 12 variables excluding error term, leaving 8 degrees of freedom.

We then construct a national average parameters, inter-municipality ranges, and improved municipality estimates using Bayesian hierarchical modeling. Because of the low number of degrees of freedom, we modify the traditional hierarchical model with a t distribution.

Elevation is excluded from the the individual municipality estimates, since it is a constant at the municipality level. However, it is included in the hierarchical step, to allow each coefficient to vary systematically with elevation (and by proxy, by species).

$$\beta_i^* \sim \mathcal{N}(\mu_\beta + \nu_\beta e_i, \tau_\beta)$$
$$E[\beta_i] =\sim t(8, \beta_i^*, SD[\beta_i])$$

4.3 Results

The results are shown in table 9. A number of features are interesting here. First, concerning the factors that affect yields, the signs of the various factors are as predicted. The effect of killing degree-days is much greater than the effect of growing degree-days, but similar to the value estimated from OLS (-3.2 here compared to -3.3 for OLS). The σ parameter, describing the range of yields, is 0.465, suggesting the the range of yields observed on plots across harvestable plants varies from 54 % to 147 % of the mean yield.

Of the factors affecting planting, autoregressivity is surprisingly high: an autoregressive factor of 0.409 suggests that only 64 % of planting is explained by the previous year's planting. This low value is partly explained by a high rate of asymptotic approach to price-driving planting.

To understand the effects of how these values change with elevation, consider the parameters predicted for the average elevation of the Arabica and Robusta species, as shown in table 10. Robusta is predicted to have a much sharper effect of killing degree-days, much larger than the average OLS estimate ($-4\,64$ compared to $-3\,3$), while Robusta has a lower effect. Similarly, Arabica has a greater effect from higher average minimum temperatures.

Arabica and Robusta are predicted to have the same costs of harvest. However, Arabica has a higher autoregressive term than Robusta, suggesting longer histories.

The transition from Arabica to Robusta represents a kind of adaptation, with warmer areas growing varieties less susceptible to high temperatures. This transition is shown in figure 11 across all of the individual municipalities. I find that cooler regions are more susceptible to both higher average minimum temperatures and killing degree days.

The average yield across years for municipalities with elevations below 800m is 1 06MT/Ha, corresponding to Robusta regions, using production divided by harvests. For regions above 800m, the average recorded yield is 1 17MT/Ha. With the integrated model, we can estimate yields using municipality-specific coefficients under partial pooling. The corresponding yields are 81MT/Ha below 800m and 1 20MT/Ha above 800m. This suggests that Robusta fields are subject to much greater harvest selection, as might be expect given the greater emphasis on yield rather than quality for Robusta.

The average prices from the World Bank Pink Sheet for Arabica over the period 1990 - 2012 are 1454USD/MT for Robusta and 2378USD/MT for Arabica. Using our estimates of yield, average incomes per hectare are 1177USD/Ha for Robusta and 2853USD/Ha for Arabica. These numbers compare well to the estimate of 1700 from section 2.2. The estimate for the cost of harvesting is 220USD/Ha for both Robusta and Arabica, and is not far from the value reported in Rodriguez and Vasquez (2009) of 434USD/Ha.

4.4 Comparison and overview

Finally, we simulate the contributing effects of each of our assumptions, applying the estimate derived above. The results are shown in figure 12. Two results are particularly significant. First, the inclusion of plant die-off at the observed levels quickly removes the age-based structure of the yields. This supports our use of a single age class for mature plants.

Each sub-figure in figure 12 provides the net present value (NPV) of the stream of yields from the perspective of year 1. The differences between these produces the other important results. The largest decrease in NPV is driven by die-off, but the largest increase from optimal harvests more than reverses its effect. Further losses of NPV are driven by delays in replanting. If those delays are random, the NPV is 5% lower, but if they are increased because of credit constraints produced by preceding lower yields, the NPV is 15% lower than without any delays.

5 Conclusion

We find that temperatures drive changes in yields, harvested areas, and planted areas. These three effects are connected by farmer decisions at the intensive and extensive margins. Extreme weather affects effect the profitability of perennials and can reduce their productive area with die-off, feeding back to the average supply of these crops. Credit constraints can not only reduce the ability of farmers to expand their fields, but they can also reduce the perceived benefit of planting those fields in the first place in light of the demand to replace damaged plants in low-yield years.

These statistical models are estimated using natural experiments, by comparing observed yields in years with different distributions of weather to estimate the effect of weather in general. These experiments completely inform our models of production. The models above include daily minimum and maximum temperatures, precipitation, and humidity. We do not include soil properties in this paper because it is impossible to do statistical experiments where soil characteristics vary over time, to see its effects. As a result, the statistical model cannot determine the effects of soil.

Some aspects of the coffee system remain internal to the models estimated here. Farmers are likely to perform activities during the season to support high yields, all they are also affected by extreme temperatures, and no attempt is made to disentangle these effects. Similarly, the drivers behind losses in yield are not explored. This *ex post* approach is both a strength and a limitation. It captures realistic relationships between weather and yields, rather than theoretical responses of the crops in an experimental setting. It can capture the environmental determinants of coffee disease spread, and their impacts implicitly. It can also be used to predict yields under climate change and weather events. However, because it cannot distinguish the social and natural causes, it makes an implicit assumption that yields will continue to respond the same way to increasing temperatures over time.

The statistical models we produce only account for about 32 - 38% of the variation in yields across time and space. The biennial cycle of coffee, for example, is not explicitly captured in our model, which considers only effects driven by weather (Bernardes et al., 2012). In addition to the biennial cycle, there are a large number of factors which drive coffee yields that are not explicitly included in this model: consumption drivers, evolving technology, changing varieties, and the governance and politics which frequently affect the coffee sector. These are all important. By limiting our analysis to the study of weather and climate change, we can better understand those elements.

Finally, these effects provide an entry point into a range of new effects from climate change. Previous studies of the effects of extreme temperatures on agriculture is directly reflected by decreases in yields. However, for perennials, the loss of full plants can be a greater threat. Farmer activities can mask the direct effects of yield losses though selective harvesting, but they are still impacted by their reduced revenue in extreme years.

References

Arak, M. V., 1967. The supply of Brazilian coffee. Ph.D. thesis, MIT.

Atkinson, C., Brennan, R., Jones, H., 2013. Declining chilling and its impact on temperate perennial crops. Environmental and Experimental Botany 91, 48–62.

Bateman, M. J., 1965. Aggregate and regional supply functions for ghanaian cocoa, 1946–1962. Journal of Farm Economics 47 (2), 384–401.

Bernardes, T., Moreira, M. A., Adami, M., Giarolla, A., Rudorff, B. F. T., 2012. Monitoring biennial bearing effect on coffee yield using modis remote sensing imagery. Remote Sensing 4 (9), 2492–2509.

Brady, M. P., Marsh, T. L., 2013. Do changes in orchard supply occur at the intensive or extensive margin of the landowner? In: 2013 Annual Meeting, August 4-6, 2013, Washington, DC. No. 150452. Agricultural and Applied Economics Association.

Devadoss, S., Luckstead, J., 2010. An analysis of apple supply response. International Journal of Production Economics 124 (1), 265–271.

Dorfman, J. H., Heien, D., 1989. The effects of uncertainty and adjustment costs on investment in the almond industry. The Review of Economics and Statistics, 263–274.

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K., Büchner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., et al., 2015. The global gridded crop model intercomparison: data and modeling protocols for phase 1 (v1. 0). Geoscientific Model Development 8 (2), 261–277.

Elnagheeb, A. H., Florkowski, W. J., 1993. Modeling perennial crop supply: an illustration from the pecan industry. Journal of Agricultural and Applied Economics 25 (01), 187–196.

Feinerman, E., Tsur, Y., 2014. Perennial crops under stochastic water supply. Agricultural Economics 45 (6), 757–766.

French, B. C., King, G. A., Minami, D. D., 1985. Planting and removal relationships for perennial crops: an application to cling peaches. American Journal of Agricultural Economics 67 (2), 215–223.

French, B. C., Matthews, J. L., 1971. A supply response model for perennial crops. American Journal of Agricultural Economics 53 (3), 478–490.

Gay, C., Estrada, F., Conde, C., Eakin, H., Villers, L., 2006. Potential impacts of climate change on agriculture: A case of study of coffee production in veracruz, mexico. Climatic Change 79 (3-4), 259–288.

Guilford, G., 2014. How climate change and a deadly fungus are threatening the world's coffee supply. Retrieved from http://www.citylab.com/weather/2014/06/how-climate-change-and-a-deadly-fungus-are-threatening-the-worlds-coffee-supply/371994/.

Guzmán Martínez, O., Jaramillo Robledo, A., Baldión Rincón, J. V., 1999. Anuario meteorologico cafetero, 1998.

Hartley, M. J., Nerlove, M., Peters, R. K., 1987. An analysis of rubber supply in sri lanka. American Journal of Agricultural Economics 69 (4), 755–761.

Hastings, D. A., Dunbar, P., 1998. Development & assessment of the global land one-km base elevation digital elevation model (globe). Group 4 (6), 218–221.

Jaramillo, J., Setamou, M., Muchugu, E., Chabi-Olaye, A., Jaramillo, A., Mukabana, J., Maina, J., Gathara, S., Borgemeister, C., 2013. Climate change or urbanization? impacts on a traditional coffee production system in east africa over the last 80 years. PloS one 8 (1), e51815.

Jordan, N., Boody, G., Broussard, W., Glover, J., Keeney, D., McCown, B., McIsaac, G., Muller, M., Murray, H., Neal, J., et al., 2007. Sustainable development of the agricultural bio-economy. Science 316 (5831), 1570.

Kalaitzandonakes, N. G., Shonkwiler, J., 1992. A state-space approach to perennial crop supply analysis. American Journal of Agricultural Economics 74 (2), 343–352.

Knapp, K. C., Konyar, K., 1991. Perennial crop supply response: a kalman filter approach. American Journal of Agricultural Economics 73 (3), 841–849.

Malkin, E., 2014. A coffee crop withers. Retrieved from http://www.nytimes.com/2014/05/06/business/international/fungus-cripples-coffee-production-across-central-america.html.

Monfreda, C., Ramankutty, N., Foley, J. A., 2008. Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global biogeochemical cycles 22 (1).

MSU AFRE Food Security (Ed.), 2015. The Evolving Role of Large and Middle Size Farms in Brazilian Agriculture. Milan, Italy.

Muschler, R. G., 2001. Shade improves coffee quality in a sub-optimal coffee-zone of costa rica. Agroforestry systems 51 (2), 131–139.

Nerlove, M., 1958. The dynamics of supply; estimation of farmer's response to price. John Hopkins University Press, Baltimore.

Nerlove, M., 1979. The dynamics of supply: retrospect and prospect. American journal of agricultural economics 61 (5), 874–888.

Pendergrast, M., 1999. Uncommon grounds: The history of coffee and how it transformed our world. Basic Books.

Price, T. J., Wetzstein, M. E., 1999. Irreversible investment decisions in perennial crops with yield and price uncertainty. Journal of Agricultural and Resource Economics, 173–185.

Rodriguez, B., Vasquez, M., 2009. Economic aspects of coffee production. In: Wintgens, J. N. (Ed.), Coffee: growing, processing, sustainable production. A guidebook for growers, processors, traders and researchers. Wiley-Vch.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., et al., 2010. The ncep climate forecast system reanalysis. Bulletin of the American Meteorological Society 91 (8), 1015–1057.

Schlenker, W., Roberts, M. J., 2009. Nonlinear temperature effects indicate severe damages to us crop yields under climate change. Proceedings of the National Academy of sciences 106 (37), 15594–15598.

Sharf, S., 2014. Mondelez to take bigger sip of $81b global coffee industry with de master joint venture. Retrieved from `http://www.forbes.com/sites/samanthasharf/2014/05/07/mondelez-to-take-bigger-sip-of-81b-global-coffee-industry-with-de-master-joint-venture/`.

Thang, T. C., 2011. Optimal investment decisions of coffee farmers in vietnam. Ph.D. thesis, The University of Western Australia.

Thurston, R. W., Morris, J., Steiman, S., 2013. Coffee: A Comprehensive Guide to the Bean, the Beverage, and the Industry. Rowman & Littlefield Publishers.

Wang, R., Alonzo, G., 2013. Foreward to the proceedings. Perennial crops for food security: Proceedings of the FAO expert workshop.

Wickens, M. R., Greenfield, J., 1973. The econometrics of agricultural supply: an application to the world coffee market. The Review of Economics and Statistics, 433–440.

Wintgens, J. N., 2009. Coffee: growing, processing, sustainable production. a guidebook for growers, processors, traders and researchers. Wiley-Vch.

Wrigley, G., 1988. Coffee. tropical agricultural series. Long man Scientific and Technical publishing: New York, 639.

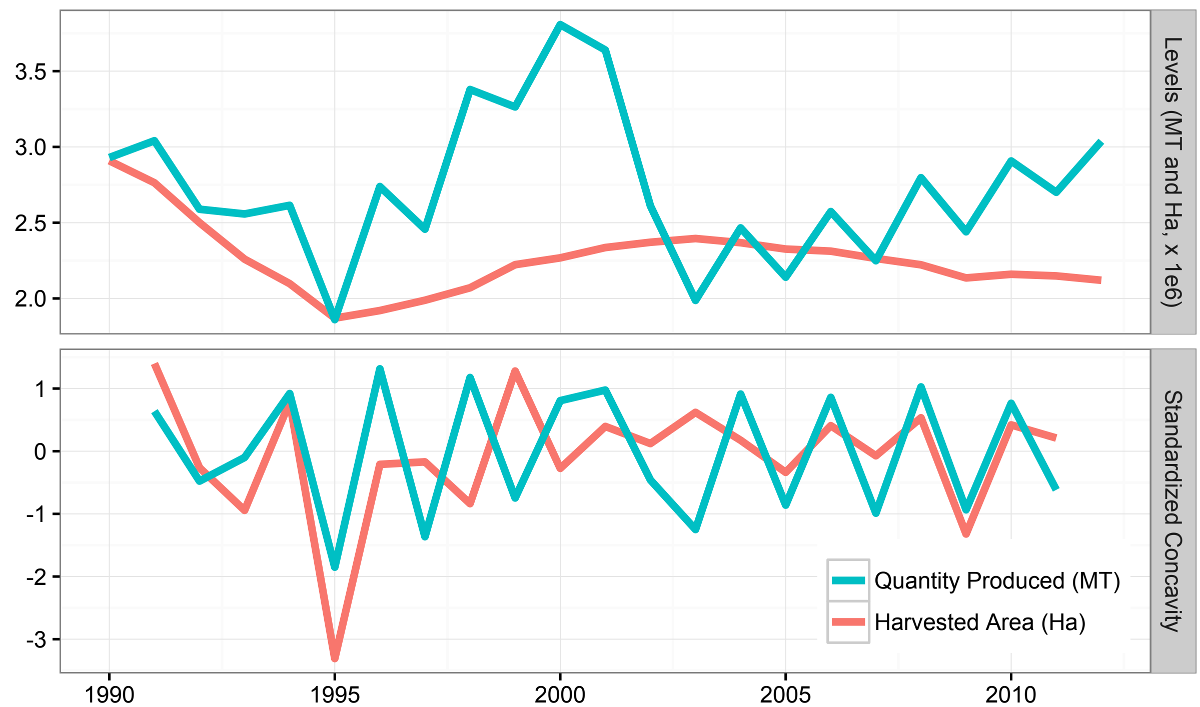

Figure 1: **Top Panel**: Absolute levels of total Brazil production (in million MT) and harvested area (in million Ha). **Bottom Panel**: The point-wise concavity of the functions $(-v_{t-1} + 2v_t - v_{t+1})$, normalized by dividing by its standard deviation. Production and harvest concavity have a correlation of 0.4.

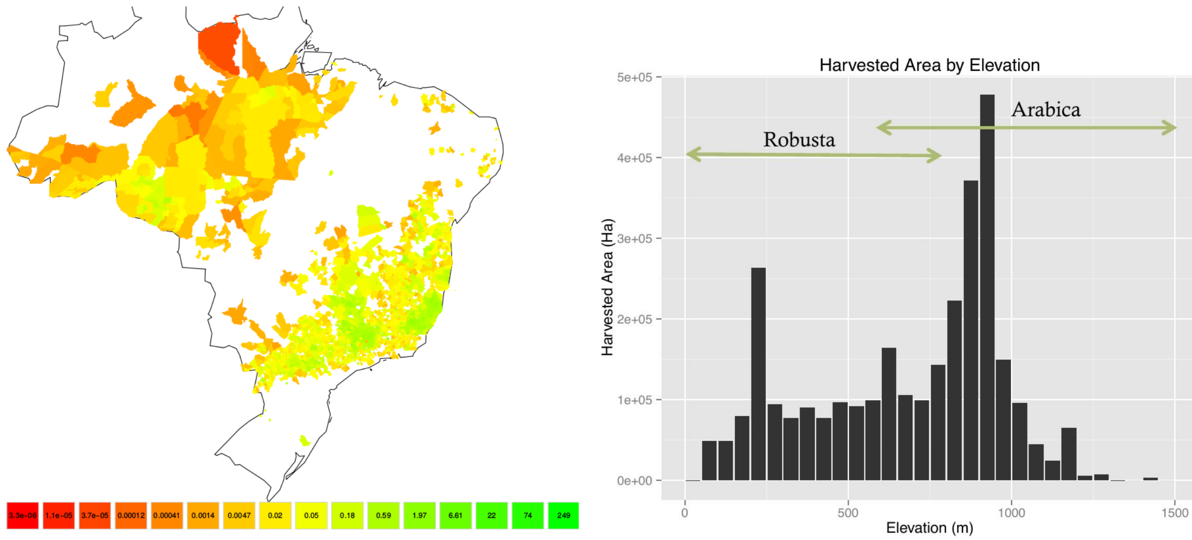

Figure 2: **Left:** Density of coffee production, as the average production divided by municipality area. Regions in green account for the majority of production. Most production occurs in the south, however there are coffee producing regions also in the southern Amazon. **Right:** Distribution of coffee producing area, displayed across the average elevation of each municipality. The greatest extent of coffee production occurs in municipalities with about 900 m of elevation, but coffee is also produced in municipalities with a much lower elevation, including a peak around 200 m. The range of typical elevations for growing Arabica and Robusta are shown above the histogram.

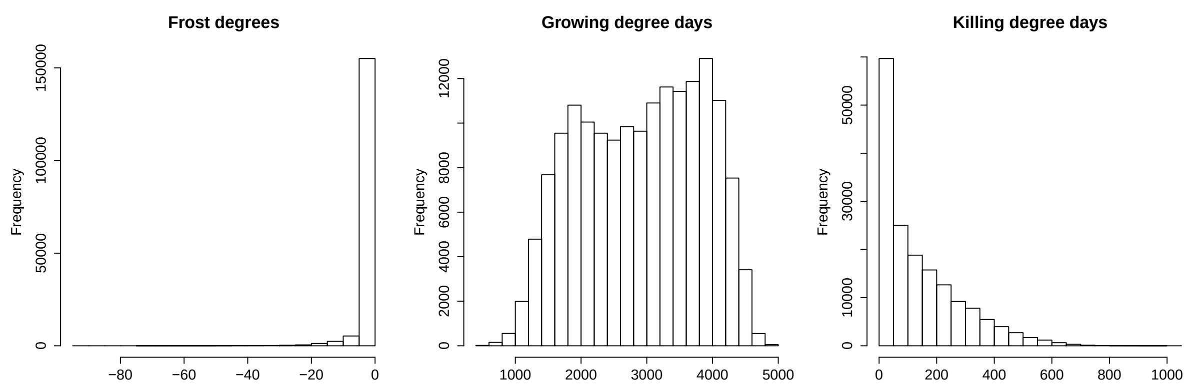

Figure 3: Histograms displaying the number of growing seasons with a given number of frost degree-days, growing degree-days, and killing degree-days. The exponential decays in frost and killing degree days are useful for capturing the impact of extreme events. The broad range of growing degree-days represented in the center histogram allows for accurate estimates of the coffee growth response.

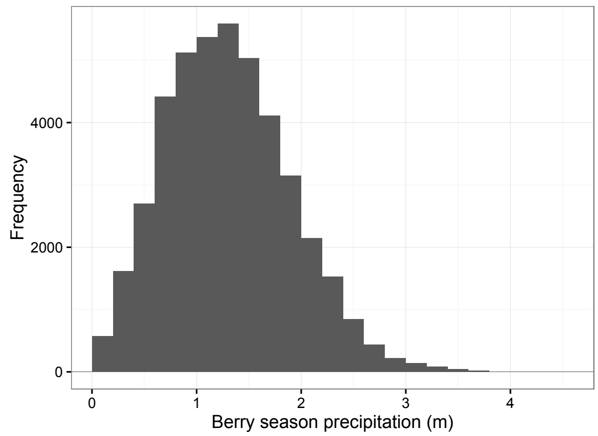

Figure 4: Histogram of the years and municipalities experiencing given total precipitation amounts, measured in meters between October 1 and September 30. The middle 50% of municipalities have total precipitation levels from 85 to 170 cm.

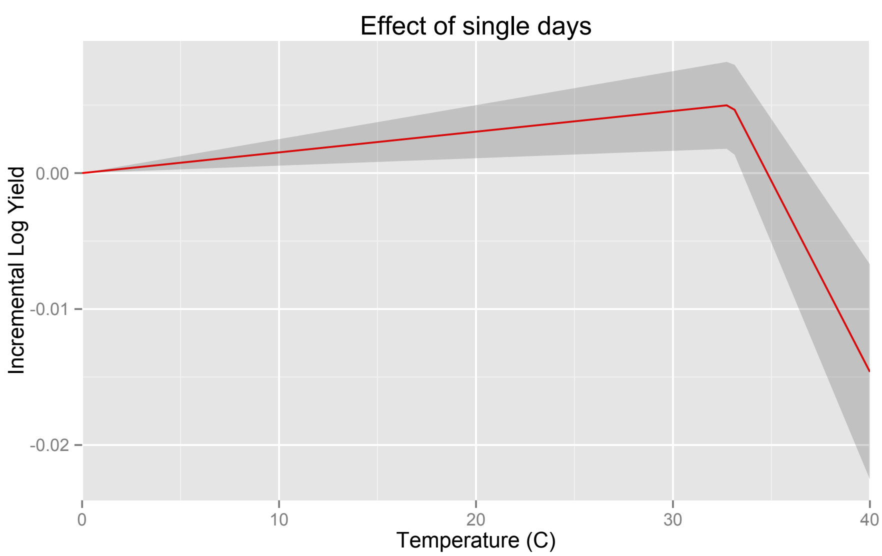

Figure 5: Marginal impact on log yields for an additional day at a given temperature. Up to 33°C, additional temperature results in greater yields. Above 33°C, this effect is sharply diminished and hot days above 35°C result in large decreases in yield. The grey band shows the 95% confidence intervals around the estimated effect for a single day at a given temperature.

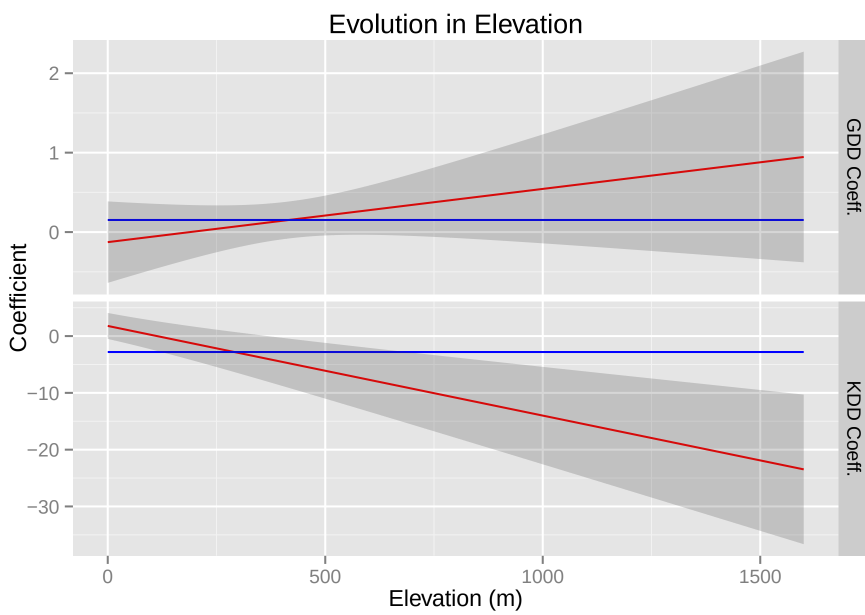

Figure 6: The effect of an additional 1000 GDD and KDD as these vary by elevation. As elevation increases, plants become more sensitive to temperatures. The effect of GDDs increases, though very slightly. The harmful effects of KDDs increase quickly.

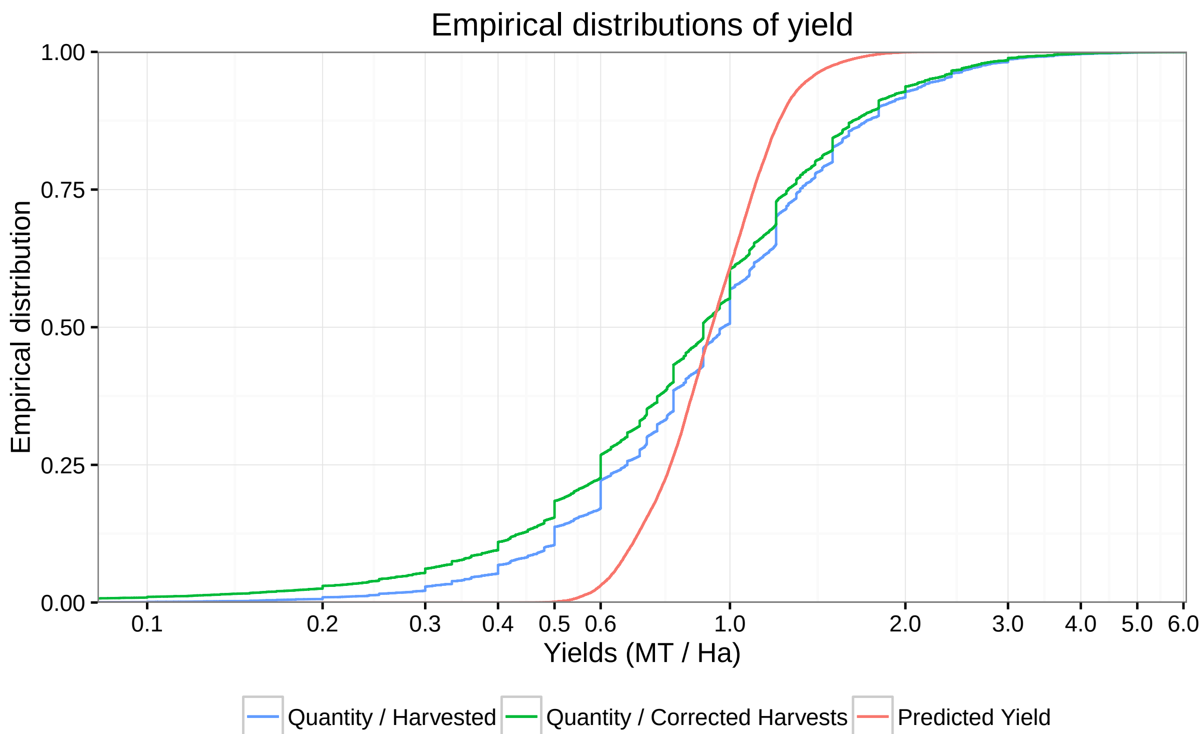

Figure 7: Empirical cumulative distributions for yield computed as q_t/h_t, yield using single-year corrected harvested areas, and yields computed by the integrated model in section 4
.

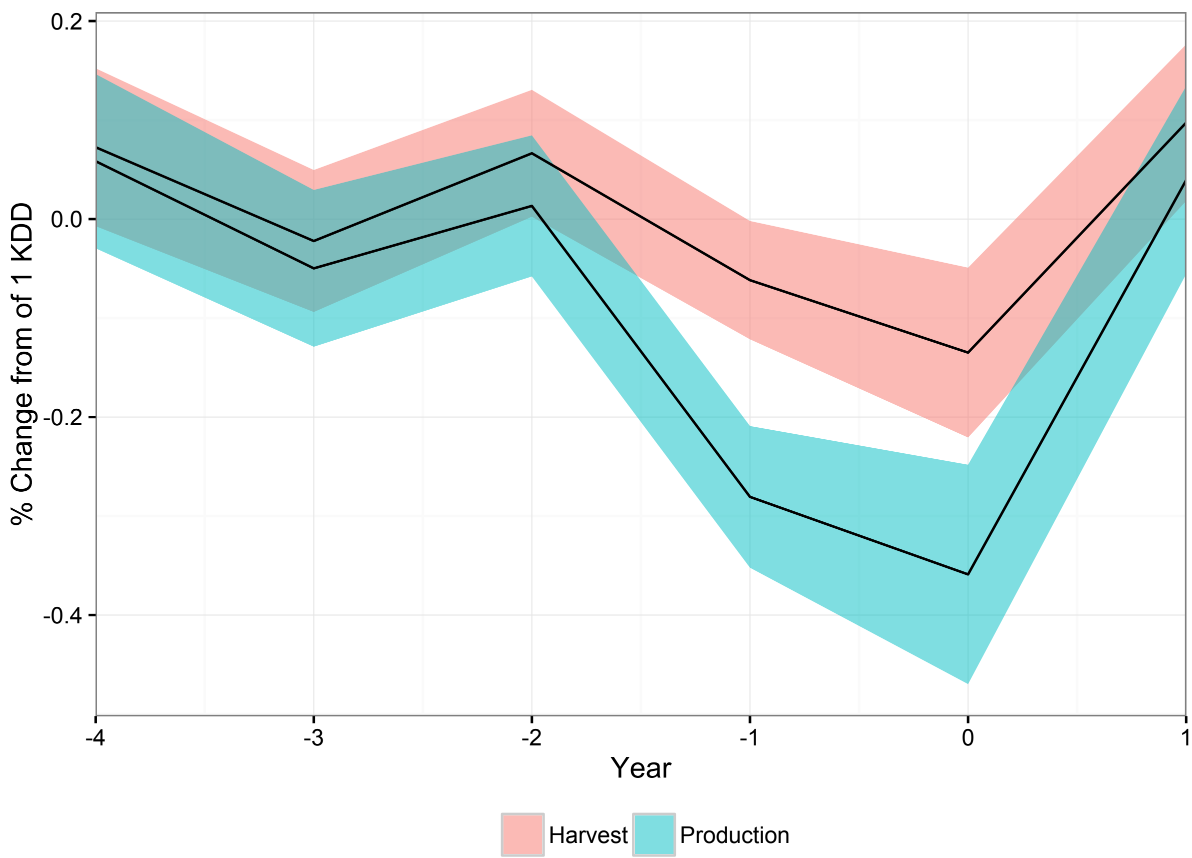

Figure 8: Coefficient on KDDs from previous years for production (blue) and harvest (red). Year -4 describes the effect of KDDs on year 0 from 4 years before it. Similarly, year 1 describes the effects from the weather in the following year, expected to be 0.

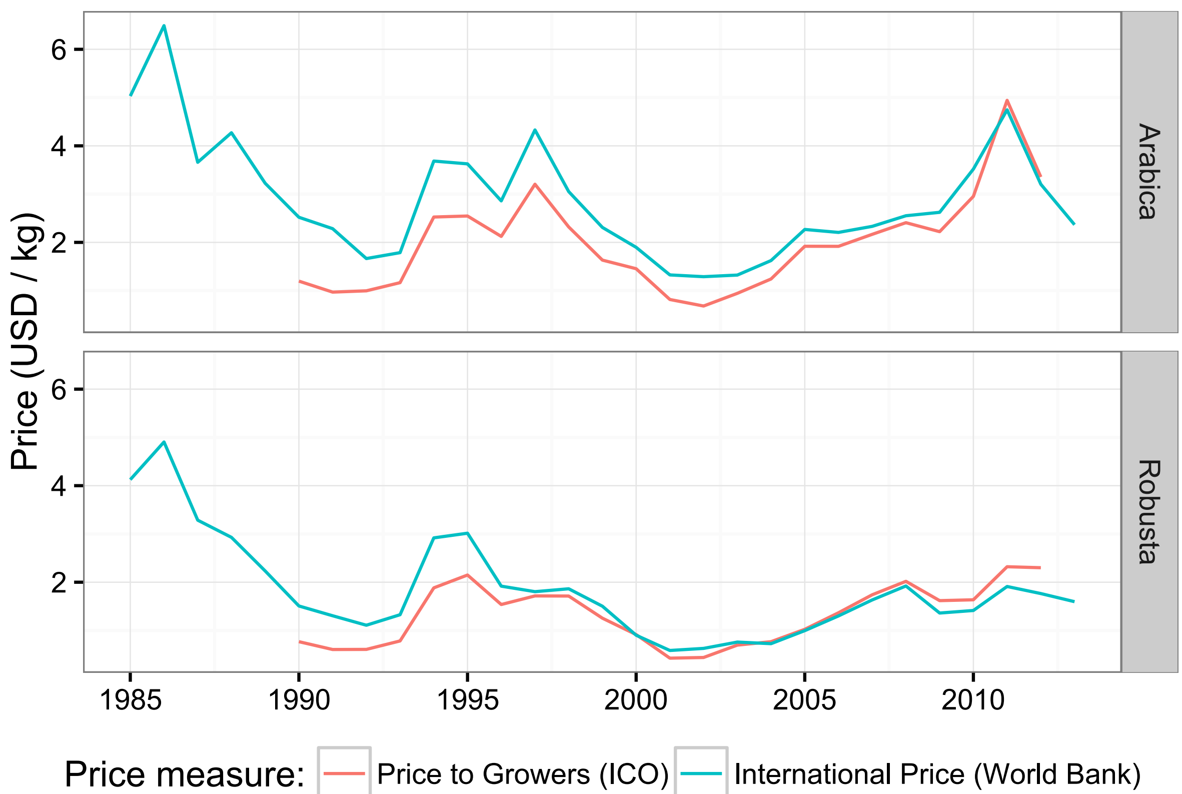

Figure 9: Prices paid to growers, according to the International Coffee Organization, and international commodity prices, according to the World Bank.

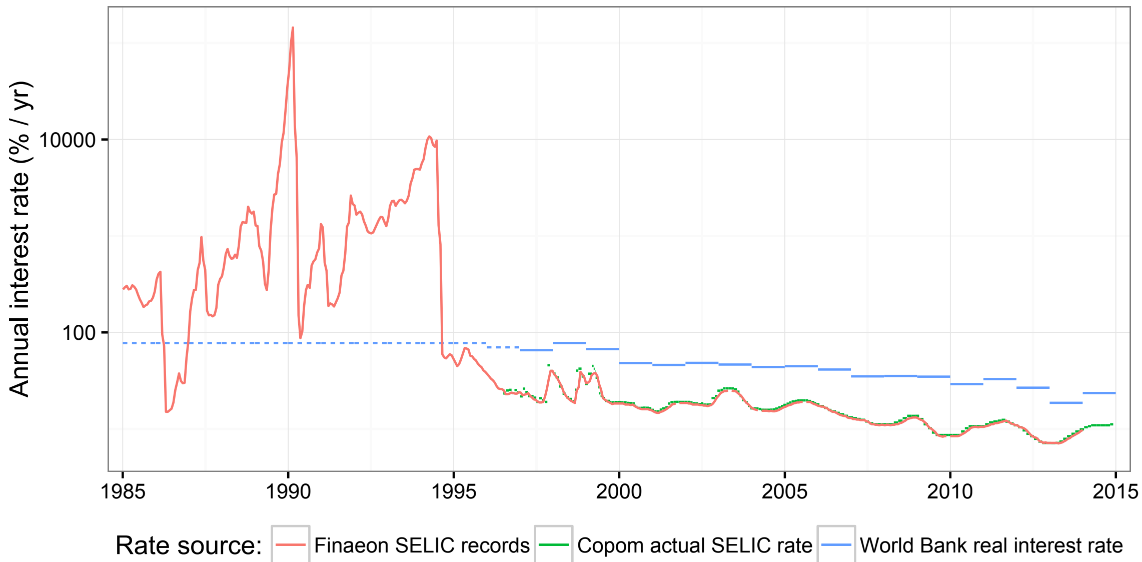

Figure 10: Brazilian interest rates from the Finaeon records, the Monetary Policy Committee (Copom) meetings, and the World Bank. Dotted lines for the World Bank prior to 1997 are from linear regression of yearly average Finaeon data on World Bank rates, capped at the highest rate observed in the World Bank history.

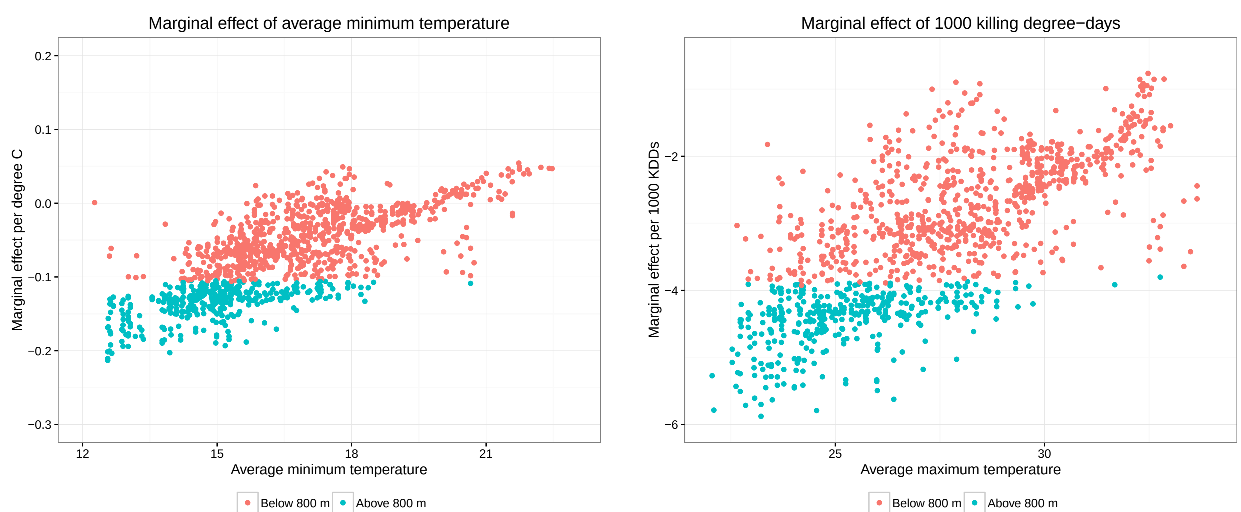

Figure 11: Estimates by municipality of the effects of average minimum temperatures vs. average minimum temperature (left), and killing degree-days vs. average maximum temperature (right). Regions are colored by their elevation, with the sharp dividing line driven by the hierarchical modeling.

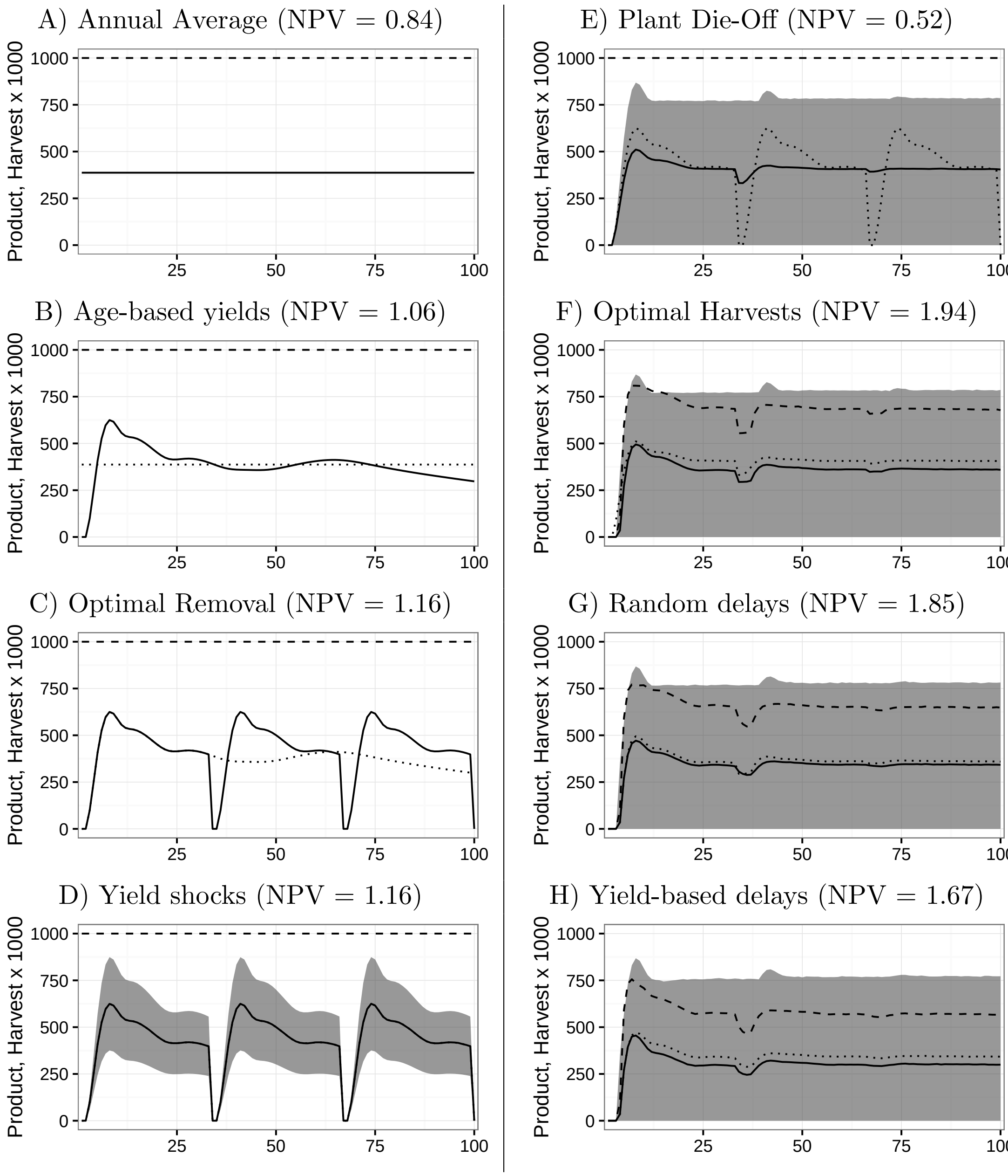

Figure 12: The incremental effects of model elements, over 100 years of planting, and the net present value at year 1. Each figure shows yields (solid line), harvest portion (dashed line, with 1000 = 100%), and the previous graph's yield (dotted line). The range shows 95% confidence intervals over Monte Carlo estimates. **(A)** An annual with the same yield the average of coffee over 100 years. **(B)** Non-stochastic yields based on ages, from Arak (1967). **(C)** Removal and replanting after the optimal number of years. **(D)** Stochastic yields, uniformly distributed between $(1-\sigma)$ and $(1+\sigma)$ of the average yield. **(E)** Random die-off of the plant, of 5% per year. **(F)** Selective harvesting, for yields $> {}^{c^H}\!/_{p}$. **(G)** Random planting delays, either after plant death or keeping plants after their optimal removal age. **(H)** Financial-driven planting delays, made more likely by prior-year yields.

Item	Small	Income	Medium	Large
Set-up	$1040		$1972	$2854
Year 1	$163		$255	$310
Year 2	$375		$580	$608
Year 3	$497	$196	$768	$608
Year 4	$493	$556	$944	$1407
Year 5	$779	$981	$1318	$1636
Year 8	$845	$981	$1362	$1845
Inputs	14%		21%	35%
Harvesting	31%		30%	32%
Labor	45%		39%	21%
Other	10%		10%	12%

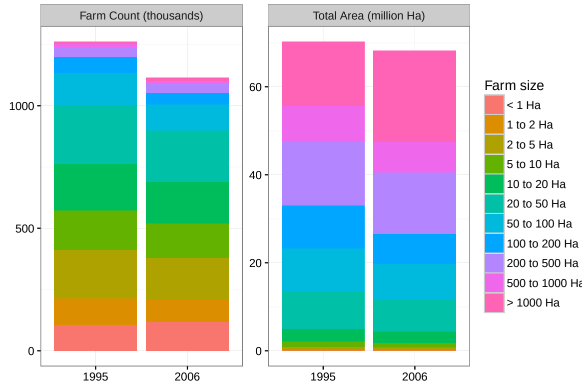

Table 1: **Left:** Cost estimates for small ($\leq$ 3 Ha), medium-sized ($>$ 3 Ha and $\leq$ 20 Ha), and large farms ($>$ 20 Ha), from Rodriguez and Vasquez (2009). Set-up costs include both activities needed for the initial establishment of a plantation, and those needed for each replanting. Other year values describe maintenance costs. Income (revenue) is shown small-farmers, and is about 60% greater for medium-sized farms, and 115% greater for large farms. **Right:** The number and total area of farms growing perennial crops in Brazil, by year (data from MSU AFRE Food Security (2015), calculations in appendix 5).

	(1)	(2)	(3)	(4)
GDDs / 1000	0.186***	−0.074	0.152***	−0.065
	(0.051)	(0.062)	(0.050)	(0.060)
KDDs / 1000	−2.243***	−2.246***	−2.806***	−3.282***
	(0.331)	(0.353)	(0.342)	(0.374)
Average Min.	−0.056***	−0.025	−0.089***	−0.022
	(0.012)	(0.016)	(0.011)	(0.015)
Precip. (m)	0.305***	0.180***	0.347***	0.164***
	(0.028)	(0.031)	(0.028)	(0.030)
Precip.^{2}	−0.364***	−0.289***	−0.366***	−0.229***
	(0.037)	(0.040)	(0.036)	(0.038)
Region FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	No	Yes
State Cubic Trends	No	No	Yes	Yes
N	43,165	43,165	43,165	43,165
R^2	0.339	0.376	0.383	0.411
Adjusted R^2	0.297	0.337	0.343	0.372
Residual Std. Error	0.553	0.538	0.535	0.523

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 2: Estimates for statistical models relating growing degree-days, killing degree-days, average minimum temperature, and precipitation to the logarithm of yields, for all Brazilian municipalities. Models differ by the form of their time controls.

	(1)	(2)	(3)	(4)
GDDs / 1000	0.527***	−0.137	0.649***	−0.037
	(0.115)	(0.158)	(0.116)	(0.152)
KDDs / 1000	−0.802	−1.975**	−1.775*	−2.219**
	(0.851)	(0.868)	(0.971)	(0.953)
Average Min.	−0.130***	0.103**	−0.187***	0.023
	(0.031)	(0.051)	(0.030)	(0.043)
Precip. (m)	0.591***	0.159*	0.520***	0.140*
	(0.089)	(0.087)	(0.090)	(0.083)
Precip.^{2}	−0.668***	−0.263*	−0.494***	−0.128
	(0.135)	(0.139)	(0.125)	(0.121)
Region FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	No	Yes
State Cubic Trends	No	No	Yes	Yes
N	3,154	3,154	3,154	3,154
R^2	0.350	0.501	0.446	0.550
Adjusted R^2	0.328	0.480	0.423	0.528
Residual Std. Error	0.409	0.360	0.379	0.343

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 3: Estimates for statistical models relating growing degree-days, killing degree-days, average minimum temperature, and precipitation to the logarithm of yields, for the top 100 Brazilian municipalities by production quantity. Models differ by the form of their time controls.

	Log yield			Log harvest hectares	
	Baseline	Linear	Indicator	Linear	Indicator
	(1)	(2)	(3)	(4)	(5)
GDDs / 1000	−0.065	−0.022	−0.179**	−0.078	−0.130
	(0.060)	(0.069)	(0.072)	(0.084)	(0.085)
KDDs / 1000	−3.282***	−4.908***	−3.128***	−2.731***	−1.818***
	(0.374)	(0.450)	(0.397)	(0.444)	(0.360)
Average Min.	−0.022	−0.018	−0.004	0.091***	0.097***
	(0.015)	(0.016)	(0.017)	(0.023)	(0.023)
Precip. (m)	0.164***	0.174***	0.159***	−0.098**	−0.104**
	(0.030)	(0.035)	(0.036)	(0.050)	(0.049)
Precip.^{2}	−0.229***	−0.269***	−0.273***	0.008	0.004
	(0.038)	(0.048)	(0.048)	(0.073)	(0.068)
Elev. x GDDs / 1000		0.117	0.394***	0.592***	0.661***
		(0.089)	(0.103)	(0.129)	(0.158)
Elev. x KDDs / 1000		−6.757***	−3.889**	−2.947***	−2.181
		(1.039)	(1.753)	(1.019)	(1.788)
Elev. x Average Min.		−0.006	−0.041*	−0.088***	−0.097***
		(0.019)	(0.023)	(0.027)	(0.032)
Elev. x Precip. (m)		0.012	0.018	0.386***	0.382***
		(0.048)	(0.058)	(0.066)	(0.073)
Elev. x Precip.^{2}		0.062	0.110	−0.306***	−0.282***
		(0.065)	(0.075)	(0.093)	(0.094)
Region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
State Cubic Trends	Yes	Yes	Yes	Yes	Yes
N	43,165	42,141	42,141	42,141	42,141
R^2	0.411	0.406	0.405	0.888	0.888
Adjusted R^2	0.372	0.367	0.366	0.880	0.880
Residual Std. Error	0.523	0.525	0.526	0.834	0.834

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 4: The effects of GDDs, KDDs, and average minimum, as each varies by elevation. All municipalities in Brazil used.

	Log yield	Log production	Log harvest
	(1)	(2)	(3)
GDDs / 1000	−0.065	0.014	0.065
	(0.060)	(0.089)	(0.072)
KDDs / 1000	−3.282***	−3.261***	−1.676***
	(0.374)	(0.460)	(0.329)
Average Min.	−0.022	0.032	0.067***
	(0.015)	(0.026)	(0.022)
Precip. (m)	0.164***	0.142***	0.002
	(0.030)	(0.047)	(0.040)
Precip.^{2}	−0.229***	−0.233***	−0.062
	(0.038)	(0.060)	(0.052)
Region FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
State Cubic Trends	Yes	Yes	Yes
N	43,165	43,165	43,165
R^2	0.411	0.870	0.888
Adjusted R^2	0.372	0.862	0.881
Residual Std. Error	0.523	0.950	0.832

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 5: The effect of weather on log yields, as the ratio of production and single-year corrected harvested areas, the effect of the same weather variables on log production and log harvests. All three decrease with extreme weather shocks.

	(1)	(2)	(3)	(4)	(5)	(6)
Harvest t-1				0.560***	0.534***	0.533***
				(0.018)	(0.021)	(0.021)
GDDs / 1000	0.065	0.142*	0.163**	0.013	0.071	0.082
	(0.072)	(0.078)	(0.079)	(0.061)	(0.067)	(0.068)
KDDs / 1000	−1.676***	−2.249***	−2.166***	−1.842***	−2.512***	−2.451***
	(0.329)	(0.372)	(0.369)	(0.334)	(0.372)	(0.370)
Average Min.	0.067***	0.063***	0.062***	0.072***	0.083***	0.083***
	(0.022)	(0.024)	(0.024)	(0.017)	(0.018)	(0.018)
Precip. (m)	0.002	−0.033	−0.026	−0.026	−0.056	−0.056
	(0.040)	(0.044)	(0.044)	(0.033)	(0.035)	(0.035)
Precip.^{2}	−0.062	−0.006	−0.024	−0.032	−0.004	−0.011
	(0.052)	(0.057)	(0.057)	(0.045)	(0.048)	(0.048)
$\Delta \log p_{t-3}$			0.297***			0.180***
			(0.057)			(0.052)
$\Delta \log p_{t-2}$		0.235***	0.305***		0.159***	0.194***
		(0.040)	(0.048)		(0.035)	(0.038)
$\Delta \log p_{t-1}$			0.275***			0.193***
			(0.061)			(0.049)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State Cubic Trends	Yes	Yes	Yes	Yes	Yes	Yes
N	43,165	35,123	35,123	36,066	29,280	29,280
R^2	0.888	0.866	0.866	0.959	0.955	0.955
Adjusted R^2	0.881	0.859	0.859	0.956	0.952	0.952
Residual Std. Error	0.832	0.849	0.848	0.502	0.492	0.491

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 6: The effects of weather and lagged prices on log changes in harvested areas, $\log {}^{h_t}/_{h_{t-1}}$.

	(1)	(2)	(3)
GDDs / 1000	0.053	0.052	0.076
	(0.063)	(0.064)	(0.068)
KDDs / 1000	-1.689^{***}	-1.505^{***}	-1.459^{***}
	(0.361)	(0.358)	(0.357)
Average Min.	0.013	0.086	0.068
	(0.014)	(0.076)	(0.076)
Precip. (m)	-0.026	-0.045	-0.049
	(0.035)	(0.035)	(0.035)
Precip.^{2}	0.046	0.039	0.049
	(0.048)	(0.048)	(0.049)
r_{t-4}			0.002^{**}
			(0.001)
r_{t-3}			-0.0002
			(0.001)
r_{t-2}		-0.003^{***}	-0.003^{***}
		(0.001)	(0.001)
r_{t-1}			0.002^{**}
			(0.001)
$\log p_{t-2}$		0.127^{***}	0.107^{***}
		(0.014)	(0.020)
Region FE	Yes	Yes	Yes
Year FE	No	No	No
State Cubic Trends	Yes	Yes	Yes
N	40,979	40,979	40,979
R^2	0.039	0.040	0.040
Adjusted R^2	-0.026	-0.026	-0.026
Residual Std. Error	1.217	1.217	1.217

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 7: The effects of weather and lagged interest rates on log changes in harvested areas, $\log h_t/h_{t-1}$.

	(1)	(2)	(3)
GDDs / 1000	0.053	0.091*	0.088
	(0.063)	(0.055)	(0.058)
KDDs / 1000	−1.689***	−1.615***	−0.878***
	(0.361)	(0.333)	(0.328)
Average Min.	0.013	0.029**	0.020*
	(0.014)	(0.012)	(0.012)
Precip. (m)	−0.026	−0.008	0.016
	(0.035)	(0.031)	(0.031)
Precip.^{2}	0.046	−0.031	−0.042
	(0.048)	(0.041)	(0.041)
r_{t-2}		−0.003***	−0.0001
		(0.001)	(0.001)
$\log p_{t-2}$		0.164***	0.366***
		(0.017)	(0.101)
KDDs(t-2) / 1000		0.459**	8.082***
		(0.211)	(1.407)
r_{t-2} KDDs(t-2) / 1000			−0.169***
			(0.030)
$\log p_{t-2}$ KDDs(t-2) / 1000			−9.715***
			(2.785)
$r_{t-2} \log p_{t-2}$			−0.004**
			(0.002)
$r_{t-2} \log p_{t-2}$ KDDs(t-2) / 1000			0.235***
			(0.049)
Region FE	Yes	Yes	Yes
Year FE	No	No	No
State Cubic Trends	Yes	Yes	Yes
N	40,979	37,666	37,666
R^2	0.039	0.015	0.016
Adjusted R^2	−0.026	−0.055	−0.055
Residual Std. Error	1.217	1.174	1.173

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 8: The effects of interactions between weather, lagged prices, and lagged interest rates on log changes in harvested areas, $\log {}^{h_t}/_{h_{t-1}}$.

Description	Var.	Average	(p-val)	Elevation	(p-val)	SD
Yield Intercept	υ	-2.16 ± 0.22	0.00	0.0024 ± 0.00068	0.00	0.62
GDD effect	γ	0.682 ± 0.086	0.00	-0.00024 ± 0.00029	0.41	0.06
KDD effect	κ	-1.33 ± 0.81	0.10	-0.0019 ± 0.0035	0.60	0.49
Average Minimum effect	μ	-0.056 ± 0.019	0.00	-7.2e-06 ± 6.83e-05	0.92	0.00
Precipitation (linear)	π	0.275 ± 0.066	0.00	-0.00042 ± 0.00023	0.06	0.11
Precipitation (quadratic)	ψ	-0.294 ± 0.096	0.00	0.00033 ± 0.00034	0.33	0.09
Range of yields	σ	0.442 ± 0.013	0.00	-5.9e-05 ± 4.5e-05	0.19	0.02
Autoregressivity	α	0.312 ± 0.0097	0.00	7.3e-05 ± 3.4e-05	0.03	0.06
Cost of harvest	c^H	0.236 ± 0.011	0.00	0.000139 ± 3.5e-05	0.00	0.06
Planting responsiveness	ϕ	0.0623 ± 0.0066	0.00	2.4e-05 ± 1.2e-05	0.05	0.01
Mortality KDD effect	δ	-0.02 ± 0.0098	0.04	1.5e-05 ± 3.3e-05	0.66	0.00
Production uncertainty	ϵ^Q	0.4098 ± 0.0075	0.00	-8.7e-05 ± 2.2e-05	0.00	0.05
Harvest uncertainty	ϵ^H	0.23 ± 0.022	0.00	-7.1e-05 ± 2e-05	0.00	0.06

Table 9: Estimates of each coefficient in the integrated model. The Average column shows the hyper-parameter across all of Brazil, at the mean elevation of coffee-growing regions. The Elevation column describes the marginal change for this parameter per meter.

Description	Variable	Robusta	Arabica
Average elevation	β	470	940
Yield Intercept	υ	-2.449	-1.319
GDD effect	γ	0.710	0.599
KDD effect	κ	-1.104	-1.987
Average Minimum effect	μ	-0.055	-0.058
Precipitation (linear)	π	0.325	0.127
Precipitation (quadratic)	ψ	-0.333	-0.179
Range of yields	σ	0.449	0.422
Autoregressivity	α	0.303	0.338
Cost of harvest	c^H	0.219	0.284
Planting responsiveness	ϕ	0.060	0.071
Mortality KDD effect	δ	-0.022	-0.015
Production uncertainty	ϵ^Q	0.420	0.379
Harvest uncertainty	ϵ^H	0.239	0.205

Table 10: The average predicted coefficients for Arabica and Robusta, as explained by their different average elevations.