The Global Allocative Efficiency of Deforestation

Prakash Mishra*

November 14, 2024

Draft: please click here for the latest version of this paper.

Abstract

This paper quantifies global inefficient and spatially misallocated agricultural deforestation: carbon emissions-intensive deforestation on land with low agricultural yields. I overcome the limitations of a reduced form descriptive analysis by incorporating spatial cost differences, agricultural trade, and cross-country non-agricultural productivity in a general equilibrium trade model to estimate how they contribute to misallocation. Against a benchmark case with a Pigouvian tax at a $190 per ton social cost of carbon, 97% of carbon emissions from deforestation since 1982 are inefficient. Strikingly, these emissions are produced by only 13% of global agricultural land. Preventing these emissions would have costed only 7% of status quo agricultural production, yielding welfare gains of $6.6 trillion since 1982. However, an equity-efficiency tradeoff results: the tax burden falls on the poorest landowners. Lastly, if countries with carbon pricing policy apply these prices to deforestation, they would deliver 5% of emissions reductions achieved under the Pigouvian benchmark.

Keywords: deforestation, carbon tax, misallocation, trade, land use, equity
JEL codes: Q56, Q57, F18, F14

*Wharton School, University of Pennsylvania. Email: mishrap@wharton.upenn.edu. I am grateful to seminar and conference participants from Wharton, Planetary Boundaries Working Group, AERE, Toulouse, and Columbia. I would additionally like to thank my dissertation committee, Arthur van Benthem, Susanna Berkouwer, Gilles Duranton, and Ulrich Doraszelski for continuous advising and feedback. I learned immensely from discussions with Jessie Handbury, Sagar Saxena, Shresth Garg, Ben Lockwood, Allan Hsiao, Anna Russo, Teevrat Garg, Mathias Reynaert, and Torben Mideksa. I am grateful to IFASTAT for proprietary fertilizer capacity and price data and Gilles Duranton for city-level metadata. This paper benefited from discussions by Seleni Cruz and Sylvain Chabé-Ferret. I gratefully acknowledge financial support from FORMAS grant number 2020-00371, the Wharton Risk Management Center's Russell Ackoff Doctoral Student Fellowship, Wharton's Global Initiative, and the Wharton Climate Center.

1 Introduction

Carbon dioxide emissions have risen precipitously over the last century, threatening to increase global temperatures by 1.5°C by 2100. Climate policy has made important strides in lowering emissions, largely by focusing on the manufacturing and transportation sectors (Gillingham and Stock 2018). However, climate policy lags in effectively combating deforestation (Balboni et al. 2023). The majority of deforestation since 1980 has cleared forests for agricultural use (Curtis et al. 2018). The resulting emissions from this agricultural deforestation drive 15% of global carbon dioxide emissions, more than the entire ground transportation (9%) and airline (4%) fleets (IPCC 2023).[1] Yet, current climate policies rarely target agricultural deforestation, often concerned about the tradeoff between climate and food security or economic development (Clausing and Wolfram 2023).[2] How can policymakers retarget deforestation to lower emissions while minimizing the loss of necessary food production?

Conceptually, a policymaker should allocate deforestation to land with high agricultural yields and low carbon emissions. Low-yield, high-emissions deforestation is thus misallocated. In theory, the efficient climate policy, a global Pigouvian tax at the social cost of carbon, corrects this misallocation.[3]

In practice, the reduction in misallocated deforestation from a Pigouvian intervention depends on three key mechanisms. First, consumers must be able to substitute away from food produced in high-deforestation regions like Brazil, potentially consuming more food produced abroad. Trade barriers hamper this substitution (Tombe 2015). Second, decreasing deforestation in a given country reduces agricultural labor demand, pushing labor into non-agricultural sectors. Non-agricultural labor productivity also varies across countries, affecting comparative advantage in agricultural production (Lagakos and Waugh 2013; Nath 2022). Finally, other, lower emissions land must have sufficient agricultural yields to meet excess food demand: this requires knowing the global, joint distribution of agricultural returns and emissions.

To measure the joint distribution of agricultural yields and emissions from deforestation, I assemble a novel, high-resolution, global dataset of yields, emissions, land use, and transportation costs. The data consists of 1.4 million 10 kilometer-by-10 kilometer plots of land. It

1. Carbon dioxide, or CO_2, emissions from deforestation come from the burning or processing of felled trees. Trees, in the process of respiration, take in carbon dioxide and release oxygen, removing carbon from the atmosphere and storing it as woody, carbon-rich, biomass.

2. Bolsonaro, controversially, on why the Amazon continues to be developed: "Indigenous people want to work, they want to produce and they can't." (Phillips 2019) A Polish official on recent deforestation: "In the West, first they built their infrastructure, and then laws to protect nature began to be introduced. We — through no fault of our own — have been developing for only 20-odd years and we are forced to (protect the environment) now, taking into account the restrictive environmental protection law." (Skiba 2023)

3. In an extension, I also consider co-benefits from biodiversity.

leverages satellite-measured tree cover from multiple sources, allowing for panel measurement of forest loss between 1982-2022. I provide three new empirical facts about deforestation between 1982 and 2022. First, 86% of deforestation-related emissions in the tropics come from agricultural clearing. Second, 10% of global land area contributes at most 13% of global yields while contributing nearly 100% of emissions. Third, among land which experienced deforestation after 1982, comparing two plots of land which started with equal vegetation, the plot with higher carbon content is deforested more. Taken together, these facts are suggestive of a misallocation: deforestation is occurring more on environmentally valuable land. This descriptive accounting, however, misses the international and multi-sector economic linkages that drives the current allocation of agricultural land.

To account for the role of prices and to simulate the effect of deforestation policies on food production, deforestation emissions, and welfare, I develop a global computable general equilibrium trade model of deforestation with a granular plot-level supply side. I employ a two-sector, two-factor framework: agriculture (using land and labor with a deforestation externality) and manufacturing (constant-returns, competitive producers using labor). Two agents enter the model: landowners, who make a discrete choice regarding land use (forest/natural use or agriculture) and workers, who supply labor either to agricultural landowners or the outside manufacturing sector.[4] Both agents consume a mix of agriculture and manufacturing goods, buying varieties of each across countries (an Armington demand system). The model is the first to nest three mechanisms which can affect the cost of reallocating agricultural deforestation: (1) plot-level heterogeneity in yields, emissions from deforestation, and transportation costs which come from the data and enters the landowners' discrete choice problem, (2) cross-country trade costs which affect consumption decisions, and (3) general equilibrium interactions across sectors from reallocating agricultural labor. I quantify the role of each mechanism in both global deforestation emissions and agricultural production.

As a key input into the model, I estimate the first global elasticity of deforestation to agricultural returns. This elasticity cannot be obtained from descriptive analysis alone. For carbon pricing interventions, it governs how landowners' land use decisions respond to changes in agricultural returns and any redistribution of surplus across landowners. For identification, I leverage variation in agricultural farm-gate prices from a large, persistent shock to demand: the accession of 26 countries to the World Trade Organization between 1995-2023. Existing WTO members trade different varieties with new WTO members, cre-

4. The static land use discrete choice problem captures the breakeven agricultural profit which generates the "first" (and most carbon intensive) deforestation event for primary forest, abstracting from secondary forest or land use dynamics.

ating differences in exposure to accession. I estimate a short-run land use elasticity of 0.09. Combining this result with data on land use shares, the average deforested plot of land earned \$36 per hectare in agricultural returns but emitted \$161 worth of CO_2 per hectare (assuming a social cost of carbon of \$190).

Given estimates of the land use elasticity, the model generates four empirical results. First, a counterfactual \$190 Pigouvian tax on carbon emissions would have abated 97% of deforestation-related emissions since 1982. This large emissions reduction requires reallocating agricultural land, lowering global agricultural production by 7%. The remaining deforestation occurs on land with high yields and low emissions. For example, in Brazil, deforestation shifts from the carbon-rich Amazonian interior to secondary forest in the Southeastern Atlantic. Agricultural production cutbacks in Brazil are further offset by increased cultivation in Chinese and Eastern African subtropical forests. Factoring in both emissions reductions and food production changes, the Pigouvian tax increases global social welfare by \$84 per ton of CO_2, or \$6.6 trillion of surplus since 1982. The magnitudes of the reallocation in emissions and production are similar across two key extensions: nonhomothetic preferences for food and endogenous irrigation investments.

Second, I quantify how plot-level heterogeneity, trade costs, and general equilibrium effects affect these results. Ignoring any one channel overpredicts the amount of emissions reductions from carbon pricing and understates losses of agricultural production. International trade costs matter most. With free trade, the same \$190/ton tax increases welfare by a further \$40/ton, nearly doubling welfare gains. The tax reduces emissions by 98.5% (a 50% decrease in remaining emissions) and global food production by 1% (with an associated 86% decrease in food costs). The welfare gain thus primarily comes from mitigating food production losses, as production more easily shifts to the US, the EU, and Canada. This supports the "greening Ricardo" hypothesis, whereby liberalizing trade improves the effectiveness of environmental policy (Le Moigne et al. 2024). Importantly, ignoring the plot-level distribution of yields and emissions and instead using country-level average yields and emissions doubles the estimated costs of the global carbon tax. Country-level data exaggerates yield losses in the EU, China, Brazil, and the US, where productive agricultural land and high-emissions deforestation are spatially distinct.

Third, the global carbon tax exacerbates inequality among agricultural landowners. The tax has concentrated costs, reallocating 10% of status quo global agricultural area to forest. Concentrated costs come from both the empirical emissions-yield distribution (high emissions landowners have low yields) and costs estimated from land use choices (high emissions landowners have high costs). I develop an analytic formula for carbon tax pass-through onto prices given an empirical distribution of landowner returns. The formula breaks down pass-

through into a "classic" ratio of supply and demand elasticities (Jenkin 1872) and a "spatial adjustment" which accounts for plot-level heterogeneity. The spatial adjustment explains 72% of the variance in carbon tax pass-through across countries. Empirically, agricultural wealth gaps arise between high-yield and high-emissions landowners within countries, rather than across countries. Within the average country, the gap between the richest 25% of landowners and poorest 25% of landowners widens by 9% under a global Pigouvian tax. Poor landowners disproportionately own high-emissions land with low returns. Wealthier landowners own high-yield, low-emissions land, and thus benefit from rising agricultural prices resulting from the carbon tax.

Fourth, I use my model to evaluate existing partial regulations which target deforestation. I focus on existing carbon price policy, active across 43 countries but rarely covering any agricultural emissions (New Zealand and, recently, Denmark are prominent exceptions; see Clausing and Wolfram 2023). Such prices are generally far below the $190 benchmark tax rate, with a (land area-weighted) average carbon price of $18 per ton. I simulate extending these prices to agriculture. Such an extension only reduces business-as-usual emissions by 5%, largely because these countries have high agricultural yields. Consequently, tax incidence is also progressive across landowners. I quantify leakage, the spillovers of emissions from regulated to unregulated regions. For every 100 tons of emissions reductions due to avoided deforestation in regulated countries, 11 are emitted through additional deforestation in unregulated countries. By coordinating across countries to target low-cost deforestation, the same damage reductions can be achieved at a 66% lower tax rate.

In sum, this paper makes three key contributions. The first is an empirical, global general equilibrium framework for measuring the food costs of abating deforestation emissions. My framework uniquely leverages plot-level microdata to correctly capture important distributional effects which highlight an equity-efficiency tradeoff in optimal deforestation policy. Second, I characterize how three key mechanisms mitigate the efficiency gains from correcting misallocated deforestation: trade, differences in comparative advantage in non-agricultural industries, and spatial costs of agricultural production. Third, I apply this framework to measure the costs of policy mistargeting and leakage. Partial regulation is more feasible than global carbon regulation, but my quantification suggests that current carbon prices would generate significant deforestation spillovers.

Related Literature. My approach brings together research on agricultural trade (Tombe 2015; Costinot, Donaldson, and Smith 2016; Sotelo 2020; Farrokhi and Pellegrina 2023; Farrokhi et al. 2023; Dominguez-Iino 2023) with research which applies discrete choice methods to land use and deforestation decisions (Scott 2013; Souza-Rodrigues 2019; Araujo, Costa, and Sant'Anna 2020; Hsiao 2022). Prior research in trade primarily highlights misallocated

agricultural production (Tombe 2015; Foster and Rosenzweig 2021; Nath 2022) or evaluates effects of trade policy on forest cover (Abman and Lundberg 2020), while I link agricultural misallocation and deforestation. Relative to prior land use estimation, particularly Araujo, Costa, and Sant'Anna (2020), I emphasize how global equilibrium spillovers can affect optimal deforestation. A significant theoretical literature motivates partial regulation-induced spillovers across multiple forested countries (Harstad and Mideksa 2017; Harstad 2020; 2023), and this emissions "leakage" has been evaluated in several environmental contexts (Fowlie 2009; Fowlie, Holland, and Mansur 2012; Hsiao 2022; Perino, Ritz, and Benthem 2023; Restrepo and Mariante 2024). My framework accounts for these equilibrium spillovers to calculate efficient forest cover.

I highlight two related papers closest to this one. First, Farrokhi et al. (2023) highlight the impact of trade liberalization on deforestation dynamics. I instead focus on how trade barriers can undercut efficiency gains from carbon taxes and thus contribute to misallocation. Second, Dominguez-Iino (2023) studies how market power in intermediary firms can affect pass-through of South American carbon taxes onto deforesting farms. In a complementary approach, I highlight theoretical and quantitatively large effects on pass-through from plot-level heterogeneity in emissions, yields, and costs (an empirical test of Head and Mayer 2023). I further refocus attention from Brazilian forest alone, allowing analysis of *which* forest basins worldwide should be targeted by policy.

My approach characterizes the distributional effects of global deforestation policy. Distributional environmental justice concerns are an increasing part of discussions of domestic place-based (Currie et al. 2015; Deryugina et al. 2019; Hausman and Stolper 2020; Christensen and Timmins 2022; Currie, Voorheis, and Walker 2023; Cain et al. 2024). I highlight interactions between inequality, environmental policy applied to deforestation, and trade (Autor, Dorn, and Hanson 2013; Borusyak and Jaravel 2024; Le Moigne et al. 2024). These equity costs have been used to motivate payments for ecosystem services in place of a carbon tax (Jayachandran 2013; Jayachandran et al. 2017), but scaling these interventions requires considering equilibrium policy impacts (Bergquist et al. 2022).

Finally, I contribute to the literature on misallocation, focusing attention on productivity dispersion *per unit emissions* (Hsieh and Klenow 2009; Syverson 2011; Bai, Jin, and Lu 2019). Recent research focuses attention on agricultural misallocation by measuring productivity dispersion across rich and poor countries (Lagakos and Waugh 2013; Foster and Rosenzweig 2021; Adamopoulos and Restuccia 2022; Gollin and Udry 2021). Other prior work assesses allocative efficiency in Australian water markets (Rafey 2023) and global groundwater extraction (Carleton, Crews, and Nath 2023).

2 Data and motivating facts.

2.1 Data

In the data, land is divided into 10×10 kilometer square grid plots. The globe is comprised of 5,040,000 such plots, of which approximately 1.3 million are non-Arctic land mass. Data consists of five major categories: land use data, productivity data, price data, and trade flows. In this section, I provide a brief overview of the key inputs to the model. I reserve other details for the data appendix C.

Land use. I use two satellite-based sources of land use data, discussed in detail in Appendix Section C.2. The first dataset from Song et al. (2018) is a continuous measurement of tree canopy cover, and primarily used for (1) descriptive analysis and (2) model validation. In this data, 220,000 plots, or approximately 15%, of global land surface experiences vegetation loss between 1982 and 2016. Not all vegetation loss is deforestation; some is degradation. I discuss this further in the data appendix.

I use European Space Agency Climate Change Initiative, hereafter ESACCI, maps of land-cover as an annual panel from 1995 to 2019 for estimation. I reclassify the base ESACCI data to calculate which pixels are forest, cropland, grassland, bare areas, and urban areas. Critically, this data allows for direct observation of *both* forest and cropland choice probabilities, where Song et al. (2018) primarily measure vegetation cover.

Deforestation emissions from the data. To calculate emissions from deforestation, I use two sets of biomass maps, both from NASA's ORNL-DAAC center. These capture stored carbon in standing vegetation in 1980 and 2010. These satellite-based measures report the amount of carbon stored in woody biomass overall.[5] I describe how I convert biomass into emissions, given two land use observations of the share of forest in 1982 and 2016 as examples s^F_{1982} and s^F_{2016}. For a plot in the data denoted ω, this is:

$$\text{emissions}(\omega; s^F_{1982}, s^F_{2016}) = \frac{44}{12}\frac{1}{2}\overbrace{\iota}^{\text{efficiency factor}}(1 - s^F_{2016})\frac{\text{biomass}(\omega)}{s^F_{1982}}$$

Breaking this down, the fractional term converts the total biomass observed on ω into a density per hectare. I then calculate the remaining biomass by assuming that biomass is constant within a plot ω; e.g., that the forest cover lost was no more or less carbon dense than the forest remaining. Next, I convert biomass into carbon using a factor of $\frac{1}{2}$, and

5. Biomass varies due to differences in growing conditions and tree species. For example, only 50% of cross-sectional biomass variation is explained by biome designations – tropical forest, subtropical forest, coniferous forest – and ecoregion designations – e.g., the Amazon, the Congo rainforest. The remaining 50% of variation is within-biome differences across trees.

carbon into carbon dioxide using a ratio of molecular weights, 44/12 (UNFCCC 2013).

Finally $\iota \in [0, 1]$ captures the fraction of potential emissions, in terms of the carbon dioxide potentially stored in a tree, that is actually emitted into the atmosphere. Concretely, for a complete burn, $\iota \approx 1$ – most woody biomass is emitted and the remaining charcoal contains a small fraction of the initial carbon. For highly efficient logging operations, $\iota \approx 0$: all of the biomass ends up in downstream timber. Unless otherwise specified, I take a pessimistic view that $\iota = 1$. I validate my emissions measure in Appendix Section C.3.1.

Potential agricultural productivity. Agricultural productivity is measured using the FAO GAEZ maps, which have a native 9-by-9 km resolution. Used extensively in prior economic research, this dataset provides information on the biophysical suitability of crops as well as information on climate, soil quality, and other environmental factors (Costinot, Donaldson, and Smith 2016). I primarily use high-input, rainfed yields for the climate epoch 1980-2010. I use a principal components analysis to reduce this large amount of potential yield data into the potential yield of a single crop index. I discuss the benefits of this approach and compare it with calorie-weighted sums of crop-level yields in Appendix Section C.3.2.

Agricultural prices. Because my agricultural productivities and the model consider a single aggregate crop output, I rely on the Food and Agriculture Organization's producer price index (PPI), an inflation-adjusted index which captures farm-gate prices at a country-year level. The producer price index is then a price index of purchases of 262 products relative to a base period of 2014-2016, when the index is set to 100:

$$p_{it} = 100 \times \sum_k \frac{p_{it}^k q_{i,2014-2016}^k}{p_{i,2014-2016}^k q_{i,2014-2016}^k}$$

where the summation index k corresponds to the set of 262 products available in the data, p^k indicate the farm-gate prices of those products, and q^k indicate production quantities. Data runs from 1991-2023 for 160 countries (an unbalanced panel).

Biodiversity data. In counterfactuals, I use data on the biodiversity value of forests. I measure the number of hectares of *potential* species habitat in each grid cell using maps of species richness from the International Union for the Conservation of Nature. These maps are at a 5 km resolution and use geographic features – elevation and temperature data, mostly – to project the number of species of amphibians, mammals, and birds in each grid cell around the world. The data also provide a rarity-weighted richness measure, which upweights species which are endemic to a given grid cell and downweights species which are suited to many other grid cells. I discuss this measure in detail in Appendix Section C.3.3.

Production and trade flows. For the estimation of the demand model and trade barriers, I use data on international production quantities and trade flows from the International Trade and Production Database for Estimation, or ITPD-E (Borchert et al. 2021; 2022). I reserve details for Appendix Section C.5.

2.2 A set of motivating facts.

The tropics are 10 times more carbon rich than the US and Europe on average. Figure 1 maps global biomass in the top panel and potential agricultural yields in the bottom panel. Biomass is extremely dense in the tropical band: the Amazon River Basin (Brazil and surrounding countries), the Congo River Basin (primarily in the Democratic Republic of the Congo), and Southeast Asia (particularly Indonesia). On average, land in the tropics (as defined by the World Wildlife Fund, see Appendix Table A.21) is 10 times more carbon-rich than land in the US or Europe. Comparing biomass with yields, there should be incentive to reallocate agricultural land away from these regions, ceteris paribus.

Agriculture is the largest source of deforestation-related emissions. Previous work has quantified the land share of deforestation due to agriculture, forestry, urban expansion, or wildfire (Curtis et al. 2018). Because wildfire among these labels is not a land use classification, I omit wildfire observations from this descriptive analysis to focus on land use drivers.[6] Combining this with my empirical emissions measure, I calculate the share of each deforestation driver in total deforestation emissions. In the tropics, 86% of emissions come from agriculture. Globally, the emissions share is 71% as logging and forestry contribute to emissions much more in non-tropical forests.

Deforestation since 1982 is positively correlated with potential carbon emissions. Focusing on land which experienced vegetation loss after 1982, I estimate the regression:

$$\text{Deforestation Rate, 1982-2016}_i = \beta_1 \log(\text{Biomass in 1980}_i) + \beta_2 \log(\text{Vegetation in 1982}_i) + v_i$$

where i is a grid plot. The coefficient β_1 describes, for two plots of land with the same starting vegetation, the differential rate of deforestation on the plot with higher potential emissions in 1982. I find a 1% increase in biomass in 1980 is associated with 0.2% more forest loss on the 180,000 grid cells which experience a loss. This correlation is statistically significant even after the inclusion of country, state, county, and even 30 km $\times$ 30 km block

6. Wildfires drove 20% of global emissions, but the authors classify land as wildfire regardless of whether land was purely left undisturbed after fire to regrow or used even temporarily as agriculture or forestry.

fixed effects. To justify such a pattern, biomass-rich forest must be better at producing agriculture than biomass-poor forest.

Deforestation is a concentrated externality. Figure 2 summarizes the previous maps, combining yield and emissions data with deforestation data in 1982-2016. 10% of global land area contributes nearly 100% of emissions from deforestation over this 40 year period. Yet, this same 10% of global land area contributes at most 13% of global agricultural yields. Among the 180,000 grid cells which experience some vegetation loss, 56% yield less in tons than they emit in tons of CO_2.

Biodiversity loss and agricultural yields from deforestation are uncorrelated. On average, deforested land after 1982 had 4 Red List tracked rare species per hectare. Within deforested land, Appendix Figure A.20 indicates no relationship between agricultural returns and the presence of rare species.

Summary. Descriptive results highlight the extreme concentration of the deforestation externality. 10% of global land area is driving 100% of deforestation emissions but only 13% of potential yields. Quizzically, deforestation rates are highest on land with higher emissions potential. However, descriptive analysis alone does not indicate which deforestation should be avoided. Prices will determine which agriculture can be reallocated at a lower cost than its environmental emissions. I next formalize this trade-off in a global, general equilibrium model.

3 Model

3.1 Economic Environment

There are of N plots of land of equal area and L workers, partitioned amongst J countries. Countries are indexed by i. Land and workers are immobile across countries. Workers are mobile within countries. Each plot of land contains a unit continuum of land parcels, each owned by a different landowner. Thus, there are two distinct agents: workers and landowners.

There are 3 goods: an agricultural good produced using land and labor, a global public good from forest-based carbon sequestration produced using land, and a competitive manufacturing good which uses labor only. Goods are imperfectly mobile across countries due to costs of trade. Consequently, prices for the two traded goods (agriculture and manufacturing) differ across countries.

Each plot of land ω is described by its land use $h \in \{F, A\}$: forest or agriculture. Plots

are endowed with a starting state: a time-invariant productivity, a share of starting forest, and a cost shifter capturing distance to nearby markets.

Agriculture and the landowners' problem. Each landowner must decide between forest and agricultural land uses. Landowners who choose agriculture earn returns π^A. Landowners require their parcel and a_i workers to produce $\eta^A(\omega)$ units of output, where $\eta^A(\omega)$ reflects the productivity of plot ω.[7] They are paid a price p_i per unit and hire labor at wage rate w_i. Plot-specific productivities $\eta^A(\omega)$ connect the model directly to yield data. Agricultural production further incurs two additional types of costs. First, plot-level costs $c^A(\omega)$ include transportation costs. Second, landowners incur a switching cost based on their parcels' starting land use: summarise the starting state of all parcels in plot ω as $s_0(\omega)$, so that the switching cost is $\phi^A(\omega; s_0(\omega))$. These switching costs include physical costs, like the cost of controlled burning of forest, and enforcement costs from existing regulation.[8] The returns to agriculture are

$$\pi^A(\omega; s_0(\omega)) = p_i \eta^A(\omega) - c^A(\omega) - \phi^A(\omega; s_0(\omega)) - a_i w_i.$$

This landowner's agricultural return is micro-founded by a Leontief production function in land parcels and labor, with total factor productivity $\eta^A(\omega)$ and a labor-specific productivity shifter a_i.

Landowners can alternatively leave their land in a natural (forested) use and earn returns π^F. Regrowing forest has a cost $\phi^F(\omega; s_0(\omega))$. By assumption, a plot which is completely forested has a zero regrowth cost. The returns to natural forest are then

$$\pi^F(\omega; s_0(\omega)) = -\phi^F(\omega; s_0(\omega)).$$

The specified structure on production gives the following discrete choice problem:

$$\pi(\omega; s_0(\omega)) \coloneqq \mathbb{E}_\epsilon \max_{h \in \{F,A\}} \pi^h_\epsilon(\omega; s_0(\omega)) = \mathbb{E}_\epsilon \max_{h \in \{F,A\}} \pi^h(\omega; s_0(\omega)) + \epsilon^h(\omega). \tag{1}$$

The shocks ϵ^A and ϵ^F are observed by the landowner but unobserved to the econometrician, representing shocks to the individual landowners' parcels within plot ω. Individual landowners allocate their full parcel to either forest or agriculture, so that an individual parcel's land use share of h, $\mu^h_\epsilon(\omega; s_0)$, is a corner solution $\mu^h_\epsilon(\omega; s_0) \in \{0, 1\}$. Then, a represen-

7. I adopt a parenthetical notation for the plot-level index ω for comparison with land use models such as Sotelo (2020): the set of plots here is discrete.

8. Conceptually, in full general equilibrium, plot-level and switching costs are paid as wages to workers in the transportation and deforestation sectors, respectively. Endogenizing these costs have vanishingly small effects on results.

tative landowners' returns $\pi(\omega; s_0)$ on plot ω is the expected return across realizations of parcel-specific shocks ϵ^h, with corresponding conditional choice probability $\mu^h(\omega; s_0)$.

Returns are determined by two endogenous prices: agricultural prices p_i and wages w_i. Agricultural goods are differentiated (Armington) across countries $i = 1, 2, \ldots, J$, so prices clear at p_i. Workers are immobile across countries, but can freely move between agriculture and manufacturing, generating the wage payment in i, w_i.

The marginal agricultural producer in a given country is a convex combination of plot-level producers. Then, in this discrete choice problem, supply curvature and producer surplus represent plot-level variation in revealed returns from deforestation, not just the distributional assumption on production shocks of each individual landowner.

Labor supply. Total labor supply in country i is perfectly inelastic at L_i. If the workers is employed in agriculture, they are paid by their landowner. All non-agricultural workers, of mass $M_i \leq L_i$, earn a wage in the manufacturing sector, which I return to below. Because workers can move freely across sectors, in equilibrium there is no wage arbitrage, so that all workers receive the same wage w_i in country i.

Trade and transportation. For the agricultural industry, I adopt a hub-and-spoke transportation network assumption (Farrokhi and Pellegrina 2023). Agricultural goods are first domestically ferried to a major city or port market, and from these "hubs" sold to an international marketplace. Within the country, the transportation cost $\tau(\omega)$ is paid by producers, a part of $c^A(\omega)$ in Equation (1). Internationally, the transportation costs are iceberg costs, T_{ij} (delivering a unit from country i to country j requires producing $T_{ij} \geq 1$ units in i). Iceberg costs thus enter final consumer prices.

I assume manufacturing locates directly in hubs. As a result, manufacturing goods only face international trade costs, the same iceberg costs T_{ij}.

The manufacturing industry. Manufacturing production is linear in labor and competitive. Manufacturing technological productivity in country i is set to $\bar{\eta}_i^M$.[9] This gives a profit maximization problem, with price of output p_i^M in country i and wage payment w_i to M_i total workers:

$$\pi_i^M(M_i) = (p_i^M \bar{\eta}_i^M - w_i)M_i$$

With an assumption of perfect competition, producers price at marginal cost. Since labor is the only input, firms in country i set prices according to:

9. Manufacturing also can emit CO_2: see Appendix A.3.

$$w_i = \bar{\eta}_i^M p_i^M \tag{2}$$

3.2 Consumption

The two agents, landowners and workers, consume agriculture and manufacturing. In principle, each individual landowner earns heterogeneous returns $\pi(\omega; s_0(\omega))$, separate from the wage payment w_i of each worker. However, my main specification adopts homothetic preferences (I relax this in later discussion) which will imply that all agents spend a constant share of their income on each good. Thus, to ease exposition, I focus on a representative consumer in this section. In country i, the representative consumer earns income Y_i.

Preferences. Representative consumers across countries face identical price elasticities σ, σ^h, discussed in detail below, but can differ in their tastes for goods based on preference shifters $\vec{\theta}$.[10] Then, in each country $j = 1, 2 \ldots, J$, a representative consumer has a nested, constant elasticity of substitution utility:

$$U_j(x) = \left(\sum_{h \in \{A,M\}} \left(\theta_j^h \sum_{i \in S} \left(\theta_{ij}^h X_{ij}^h \right)^{\frac{\sigma^h - 1}{\sigma^h}} \right)^{\frac{\sigma^h(\sigma-1)}{\sigma(\sigma^h-1)}} \right)^{\frac{\sigma}{\sigma-1}} - \mu D \tag{3}$$

In the outer nest, they divide their income between agricultural and manufacturing goods, indexed $h \in \{A, M\}$. For each sector, the inner nest decision allocates expenditure across countries.[11] The outer elasticity $\sigma \in (0, \infty)$ will dictate the relative complementarity or substitutability across agricultural and manufacturing consumption nests. Aggregate quality differences in agricultural production enter through differences in θ_i^A.

Conditional on their outer nest demand, consumers decide how to allocate demand across countries' differentiated varieties. In this inner nest, cross-country substitutability of varieties is governed by parameter $\sigma^h \in (1, \infty)$: higher values of the elasticity will indicate greater indifference between countries of origin.

Consumers experience a disutility from total global carbon emissions D, a global public bad, at a utility cost μ. Because there are many consumers, the public goods problem prevents them from internalizing damages D from deforestation. I discuss this further in the planner's problem description in Appendix A.3.

10. I later partially relax the assumption that consumers are identically price elastic. This assumption is made for tractability as consumers will demand goods across over 100 countries.

11. The Armington assumption (whereby each country produces a differentiated variety) imposes non-zero trade across pairs of countries where there is a finite trade cost.

Budget constraint. To complete the utility maximization problem, I define an aggregate budget constraint. Expenditure on a given variety (country) i of agriculture is $E_{ij}^A = (1 + T_{ij})p_i X_{ij}^A$; in manufacturing, it is $E_{ij}^M = (1 + T_{ij})p_i^M X_{ij}^M$. The representative consumer faces an aggregate budget constraint on expenditures:

$$\sum_{i=1}^{J}(E_{ij}^A + E_{ij}^M) = Y_j.$$

With a utility function and budget constraint defined, I next derive expenditure shares for the representative consumer in each nest of the utility maximization problem. I reserve the full derivation for Appendix Section A.5.

Inner nest expenditures. I begin with results for inner nest expenditures, which hold fixed the outer nest expenditures E_j^h in a given sector h. In the manufacturing industry, the representative consumer spends share $\lambda_j^{i|M} = \frac{E_{ij}^M}{E_j^M}$ of total manufacturing expenditures E_j^M on goods from i,

$$\lambda_j^{i|M} = \frac{(\theta_{ij}^M \bar{\eta}_i^M)^{\sigma^M - 1}(T_{ij}w_i)^{1-\sigma^M}}{\sum_{n=1}^{J}(\theta_{nj}^M \bar{\eta}_n^M)^{\sigma^M - 1}(T_{nj}w_n)^{1-\sigma^M}}. \tag{4}$$

Similarly, the representative consumer spends share $\lambda_j^{i|A} = \frac{E_{ij}^A}{E_j^A}$ of total agricultural expenditures E_j^A on goods from i:

$$\lambda_j^{i|A} = \frac{(\theta_{ij}^A)^{\sigma^A - 1}(T_{ij}p_i)^{1-\sigma^A}}{\sum_{n=1}^{J}(\theta_{nj}^A)^{\sigma^A - 1}(T_{nj}p_n)^{1-\sigma^A}}.$$

Inner nest price indices. Using these inner nest expenditure shares, I derive the Dixit-Stiglitz price index of each sector P_j^h. Intuitively, these are price indices for the complete basket of food or manufacturing items. In manufacturing, the price index is

$$P_j^M = \left(\sum_{n=1}^{J}(\theta_{nj}^M \bar{\eta}_n^M)^{\sigma^M - 1}(T_{nj}w_n)^{1-\sigma^M}\right)^{\frac{1}{1-\sigma^M}}.$$

Similarly, in agriculture, the price index is

$$P_j^A = \left(\sum_{n=1}^{J} (\theta_{nj}^A)^{\sigma^A - 1} (T_{nj} p_n)^{1-\sigma^A} \right)^{\frac{1}{1-\sigma^A}}.$$

Outer nest expenditures. Finally, moving to the outer nest, I derive expenditure shares on each sector λ_j^h as a function of these price indices for each sector.

$$\lambda_i^h = \begin{cases} \frac{(\theta_i^h)^{\sigma-1}(P_i^h)^{1-\sigma}}{\sum_{h \in \{A,M\}} (\theta_i^h)^{\sigma-1}(P_i^h)^{1-\sigma}} & \text{if } \sigma \in (0,1) \cap (1,\infty) \\ \theta_i^h & \text{if } \sigma = 1 \end{cases} \tag{5}$$

where I single out the Cobb-Douglas case when $sigma = 1$. Thus, as a share of aggregate income in i, Y_i (defined below), the representative consumer spends $E_{ji}^A = \lambda_i^A \lambda_i^{j|A} Y_i$ on agricultural goods from j. I will use the notation $\lambda_{ji}^h = \lambda_i^h \lambda_i^{j|h}$ to denote the unconditional share of income that the representative consumer in i spends on good h from country j.[12]

3.3 Equilibrium: Business-as-usual

I describe here aggregation of supply and demand curves and the market equilibrium concept. I refer to this equilibrium as the business-as-usual (BAU) equilibrium.

Aggregation. For plot ω, the expected plot-level supply curve is $q(\omega) = \eta^A(\omega)\mu^A(\omega; s_0(\omega))$, where $\mu^A(\omega; s_0(\omega))$ represents the average share of land across individual parcels within plot ω devoted to agriculture. Aggregating over the set Ω_i of N_i plots ω in i gives an aggregate supply curve:

$$Q_i^A = \sum_{\omega=1}^{N_i} \eta^A(\omega)\mu^A(\omega; s_0(\omega))$$

I next aggregate over landowners' rents to get an aggregate rental rate:

$$\Pi_i = \sum_{\omega=1}^{N_i} \sum_{h \in \{F,A\}} \mathbb{E}_\epsilon \pi_\epsilon^h(\omega; s_0(\omega))$$

Aggregate rents depend on the functional form of the distribution of parcel-level shocks

12. Appendix A.6 provides comparative statics for the expenditure share λ_{ij}^h, including a discussion of income effects.

per plot ω, $\epsilon^h(\omega)$. For Type-I extreme value shocks, the aggregate rent takes the following closed-form solution:

$$\Pi_i = \frac{1}{\gamma} \sum_{\omega=1}^{N_i} \left[\aleph + \log \sum_{h \in \{F,A\}} \exp[\gamma \pi^h(\omega; s_0(\omega))] \right],$$

where $\aleph \approx 0.557$ is the Euler-Mascheroni constant and γ is .... The formal derivation is provided in Appendix A.1.

Define total income Y_i in the economy of i as the sum of wage income and landowner rents.

$$Y_i = w_i L_i + \Pi_i.$$

Finally, define aggregate manufacturing labor demand M_i as the residual labor supply after agricultural landowners make their land use decisions[13]:

$$M_i = \left(L_i - a_i \sum_{\omega=1}^{N_i} \mu^A(\omega; s_0(\omega)) \right).$$

Market clearing. The equilibrium must identify $2J$ endogenous prices: J country-level agricultural prices p_i^A and J country-level wages w_i. Goods market clearing[14] in agriculture equates supply (from the landowners' problem) to international expenditures (from the consumer choice problem), giving J unique equilibium conditions:[15]

$$\sum_{j=1}^{J} E_{ij}^A = p_i Q_i^A. \tag{6}$$

Next I construct equilibrium wages. The manufacturing industry absorbs residual labor supply net of labor used by agriculture. Additionally, all manufacturing surplus is paid out to manufacturing laborers, providing a balance condition:

13. The elasticity of the residual manufacturing labor supply in i is pinned down by the wage elasticity of land use decisions μ^A and a_i. Thus, by assumption, manufacturing labors supply is more wage elastic in countries with higher share of labor in agricultural value added (smallholders).

14. Using goods market clearing to close the model implies that price will equal the marginal cost of the last producing agriculturalist. Analogous to models of electricity markets, firms are ordered on their heterogeneous marginal costs (the so-called merit order).

15. Prices do not require a normalization because forest returns are already normalized to be 0.

$$\sum_{j=1}^{J} E_{ij}^{M} = w_i M_i \tag{7}$$

I normalize one wage to 1, as Walras's Law applies to the manufacturing industry. Existence of the equilibrium follows from an application of Brouwer's theorem, formally derived in Appendix A.2.[16]

Summarizing the approach. My model modifies a classic Armington model of agricultural trade to allow for a richer, disaggregated supply-side. This richness is required to understand a carbon tax which is measured in levels: the dollar cost of carbon needs to be weighed against the dollar returns to agriculture, challenging the usual identification of trade models in relative terms (Arkolakis, Costinot, and Rodríguez-Clare 2012). In this way, my model is closest to Dominguez-Iino (2023). However, I assume perfectly competitive firms, where he assumes a richer downstream supply chain with market power. Instead, I add granularity to the supply-side farmers' decision. In the next two sections, I will demonstrate how this granularity matters for counterfactuals of interest.

3.4 Internalizing carbon damages from deforestation

Damages from deforestation. Production of agriculture has an unpriced secondary product: carbon emissions. Deforestation-related carbon emissions are heterogeneous. Connecting the model directly to empirical biomass, emissions from deforesting plot ω are $d(\omega)$. Total global emissions globally sum over plot-level emissions,

$$D = \sum_{\omega=1}^{N} d(\omega)\mu^{A}(\omega; s_0). \tag{8}$$

Counterfactual: Pigouvian tax. I consider the effect of a global Pigouvian tax t on deforestation whereby a global entity taxes agriculturalists at the rate of the marginal damage of carbon, $t = t^{SCC}$, for each ton of emissions.[17] Under such a Pigouvian tax, an agriculturalist acts as though they face higher switching costs from forest to agriculture.

$$\pi_{\epsilon}^{A}(\omega; s_0(\omega), t^{SCC}) = \pi_{\epsilon}^{A}(\omega; s_0(\omega), 0) - t^{SCC} d(\omega)$$

16. Numerically, my model converges to a unique equilibrium across a wide range of starting guesses.
17. I discuss the Pigouvian tax as the solution to a planner's problem in Appendix A.3.

I assume that the tax revenue $\mathcal{T}_i = \sum_{\omega=1}^{N_i} t^{SCC} d(\omega) \mu^A(\omega; s_0(\omega))$ is distributed as a lump-sum, so the economy-wide income, including tax revenues, is $Y_i = \Pi_i + w_i L_i + \mathcal{T}_i$.

Identification. I demonstrate partial identification of the Pigouvian tax counterfactual in changes – the so-called exact hat approach – in Appendix Section A.4 (Dekle, Eaton, and Kortum 2007). Identification in changes is only partial because the supply-side model requires information on the *level* of starting prices and wages in order to calculate the impact of a carbon tax in levels (Kalouptsidi, Scott, and Souza-Rodrigues 2021). With data on starting prices, trade flows, agricultural productivity $\eta^A(\omega)$, and emissions $d(\omega)$, the counterfactual is fully identified under a parametric assumption on the error distribution ϵ^h from the landowner's problem in Equation (1) and the demand elasticities, $\vec{\sigma}$.

Measuring welfare implications of counterfactuals. Welfare in business-as-usual is measured by real income (see, e.g., Arkolakis, Costinot, and Rodríguez-Clare 2012; Shapiro 2016).

$$W_i = \frac{Y_i}{P_i^C}$$

To measure *changes* in welfare due to a counterfactual tax policy t, I need to separately account for changes in the welfare of landowners (identified as changes in levels) and workers (identified in relative changes). I use equivalent variation to quantify welfare in changes. Using my results regarding model identification, equivalent variation for landowners and workers can each be written as changes in real income due to a tax t:

$$EV_i^{LAND}(\omega; t) = \frac{1}{\gamma}\left[\frac{\log\left(\sum_{h\in\{F,A\}} \exp[\gamma\pi'(\omega; s_0(\omega), t)]\right)}{P_i^C(t)} - \frac{\log\left(\sum_{h\in\{F,A\}} \exp \gamma\pi'(\omega; s_0(\omega), 0)]\right)}{P_i^C(0)}\right]$$

$$EV_i^{WORK}(t) = \frac{w(0)L_i}{P_i^C(0)}\left[\frac{w(t)}{P_i^C(t)}\frac{P_i^C(0)}{w(0)} - 1\right]$$

Cross-country distributional differences in landowner incidence are pinned by changes in the real price of consumption, while within-country differences depend on the landowners' problem. Aggregate welfare costs are the weighted sum of equivalent variation across landowners and workers, with aggregation weights given by business-as-usual welfare shares $\alpha_i^{WORK} = 1 - \alpha_i^{LAND} = \frac{\frac{w_i}{P_i^C}}{W_i}$.

$$EV_i(t) = \alpha_i^{WORK} EV_i^{WORK}(t) + \alpha_i^{LAND} \sum_{i=1}^{N_i} EV_i^{LAND}(\omega; t) \tag{9}$$

Summary of counterfactual. The model thus identifies counterfactual prices p_i, wages w_i, food production Q_i, emissions D, and welfare $EV_i(t)$ under a level carbon tax t.

3.5 Pass-through of carbon taxes

The key difference between the model I present and contemporary trade models is the use of a disaggregated, plot ω-level joint distribution of productivity $\eta^A(\omega)$ and emissions $d(\omega)$. In contemporary work, parametric assumptions on these distributions allows analytic characterizations of welfare effects of trade policy. I diverge for two reasons. First, I analyze a carbon tax, which operates in levels, as opposed to changes in trade barriers T_{ij}. Second, I argue that the richness of the plot-level joint distribution matters for final results. I make an analytic case for this second claim in this section.

I first trace out the intuition. Agricultural prices clear at the country-level. Then, rich subnational heterogeneity allows for arbitrary wedges between producers on the margin, who are driving prices, and producers further up the supply curve, who drive the average costs of the policy. These differences can matter if, for example, much of the policy action on the margin does not affect firms who are inframarginal, creating large divergence between these producers. This section explores these dynamics through the pass-through of a carbon tax. The key formula will consist of a classic pass-through result which is "spatially modified" to account for differences in the joint distribution of yields and emissions across markets.

For expositional clarity, in this section, I derive a partial equilibrium pass-through formula *holding wages fixed*. To begin, I totally differentiate the agricultural market clearing condition from Equation (6) under a tax t:

$$Q_i^A = \sum_{j=1}^{J} (1 + T_{ij}) X_{ij}^A \implies \frac{dp_i}{dt} = \frac{-\frac{dQ_i^A}{dt}}{\frac{dQ_i^A}{dp_i} - \sum_{j=1}^{J} \frac{d(1+T_{ij})X_{ij}^A}{dp_i}}$$

The above price pass-through rate resembles a classic ratio of supply and demand elasticities. The change in quantities with respect to a tax in the numerator $\frac{dQ_i^A}{dt}$ is different from – though related to – the supply derivative in the denominator $\frac{dQ_i^A}{dp_i}$. Breaking $\frac{dQ_i^A}{dt}$ down,

$$\frac{dQ_i^A}{dt} = \sum_{\omega=1}^{N_i} \frac{d}{dt} q(\omega) = \sum_{\omega=1}^{N_i} \eta^A(\omega) \frac{d\mu^A(\omega)}{dt} = -\sum_{\omega=1}^{N_i} \eta^A(\omega) \tilde{s}(\omega) [\gamma d(\omega)] \coloneqq -\gamma N_i \mathbb{E}[d(\omega), \eta^A(\omega) \tilde{s}(\omega)]$$

where $\tilde{s}(\omega) = \frac{d\mu^A(\omega)}{d\pi^A(\omega)}$ is the semi-elasticity of agricultural land use shares to expected returns

(e.g., for an i.i.d. logit error term $\epsilon^h(\omega)$, this is $\mu^A(\omega; s_0(\omega))[1-\mu^A(\omega; s_0(\omega)]$). Define $\mathbb{E}[XY]$, the uncentered covariance among realizations of (X, Y) in Ω_i.[18] Then,

$$\frac{dQ_i^A}{dt}\left(\frac{dQ_i^A}{dp}\right)^{-1} = -\frac{\mathbb{E}_i[d(\omega), \eta^A(\omega)\tilde{s}(\omega)]}{\mathbb{E}_i[\eta^A(\omega), \eta^A(\omega)\tilde{s}(\omega)]}$$

Plugging this identity back into the initial expression for $\frac{dp}{dt}$ derived from market clearing returns a "spatially modified" aggregate pass-through formula. Define elasticities of supply $S_i^{A,p} = \frac{d\log Q_i^A}{d\log p_i}$ and demand $D_i^{A,p} = \frac{d\log \sum_{j=1}^{J}(1+T_{ij})X_{ij}^A}{d\log p_i}$. Then, pass-through is:

$$\frac{dp_i}{dt} = \overbrace{\frac{\mathbb{E}_i[d(\omega), \eta^A(\omega)\tilde{s}(\omega)]}{\mathbb{E}_i[\eta^A(\omega), \eta^A(\omega)\tilde{s}(\omega)]}}^{\text{adjustment for spatial heterogeneity}} \underbrace{\frac{1}{1-\frac{D_i^{A,p}}{S_i^{A,p}}}}_{\text{classic Jenkin (1872) pass-through}} \tag{10}$$

This pass-through formula encapsulates the three key empirical objects required to estimate the impact of a global carbon tax. First, my data pins down yields $\eta^A(\omega)$ and emissions $d(\omega)$. Second, I estimate the price elasticity of agricultural production, $S_i^{A,p}$. Third, pass-through also depends on the elasticity of demand in agriculture. While this parameter has been extensively estimated in prior work, I provide an estimate of this parameter as well. Two of my model's three mechanisms show up in this partial equilibrium pass-through: plot-level heterogeneity in yields, costs, and emissions all matter for the spatial heterogeneity term and the elasticity of supply $S_i^{A,p}$, while trade costs show up in the demand elasticity $D_i^{A,p}$.

In equilibrium, wages also vary in response to the Pigouvian tax. In general equilibrium, then, deadweight loss is shared between workers' wages and landowners' returns. In Appendix Section A.7, I derive the general equilibrium pass-through expressions for prices and wages. The final general equilibrium pass-through rate is a linear combination of Equation (10) and the equivalent pass-through rate when prices are held fixed and wages respond to the tax. General equilibrium predictions additionally require estimates of the cross-price elasticity of demand between agriculture and manufacturing, σ, and σ^M, the cross-country elasticity of substitution in the manufacturing sector.

18. If firm production is endogenous, pass-through depends on the distribution of firm-level supply elasticities, $\frac{dq(\omega)}{dp}$. I consider one such extension in Appendix Section D.9. Ongoing work considers a plot-level production function, extending (Sotelo 2020) with methods from (Doraszelski and Jaumandreu 2018), though final results are fairly robust to perturbing $\eta^A(\omega)$.

4 Empirical strategy

The model clarifies the key ingredients involved in describing the returns from deforestation. I estimate three key parameters: a land use elasticity, from Equation (1), and two cross-country elasticities of substitution σ^A, σ^M, each from sectoral trade shares in Equation (A.1). The land use elasticity governs the cost in dollars required to produce a change in landowners' deforestation decisions. The elasticities of substitution determine consumers' willingness to substitute domestic production with production abroad in response to price increases.

4.1 Landowner's problem: from the theory to the data

Deriving a regression equation. The focal estimation on the supply-side concerns the landowner's problem, Equation (1). To derive a regression equation, I assume that the error terms $\epsilon^F(\omega), \epsilon^A(\omega)$ are drawn from two independent and identical Type I generalized extreme value distributions.[19] This assumption implies a closed-form conditional choice probability for each land use, $\mu^F(\omega; s_0(\omega))$ for forested land $h = F$ and $\mu^A(\omega; s_0(\omega))$ for agricultural land $h = A$. Choice probabilities are:

$$\mu^h(\omega; s_0(\omega)) = \frac{\exp[\gamma \pi^h(\omega; s_0(\omega))]}{\sum_{k \in \{F,A\}} \exp[\gamma \pi^k(\omega; s_0(\omega))]}$$

where the parameter γ is the reciprocal of the scale parameter of the type-I extreme value $\epsilon^h(\omega)$. Next, I derive a linear regression from nonlinear choice probabilities (Hotz and Miller 1993). Taking logs of the above choice probability, and differencing across land uses gives:

$$Y(\omega) \coloneqq \log \mu^A(\omega; s_0(\omega)) - \log \mu^F(\omega; s_0(\omega)) = \gamma[\pi^A(\omega; s_0(\omega)) - \pi^F(\omega; s_0(\omega))]$$

I allow for measurement error through a statistical residual $\epsilon^{FA}(\omega)$. I introduce an additional non-structural error term $\xi(\omega)$ which represents unobservable returns to land in the data but outside my model. The structural error will represent a source of potential endogeneity in my later discussion of threats to identification. I obtain a regression:

19. This assumption limits attention to land with interior solutions: see Appendix Table A.22 for a sample selection balance table. Interior land has higher biomass, yield, population, and market access than excluded areas. Conceptually, 10×10 km grid cells are still large relative to the average farm – which is smaller than 5 ha, or 0.005% of grid cell area (Foster and Rosenzweig 2021) Thus, ϵ should be interpreted as idiosyncratic yield shocks across many very small farms.

$$Y(\omega) = \gamma[\pi^A(\omega; s_0(\omega)) - \pi^F(\omega; s_0(\omega)) + \xi(\omega)] + \epsilon^{FA}(\omega)$$

As indicated by the above regression equation, γ can also be interpreted as a deforestation semi-elasticity. It indicates the percent increase in the ratio of agricultural land to forested land in response to a \$1 increase in relative agricultural revenue $\pi^A(\omega) - \pi^F(\omega) + \xi(\omega)$.

Outside option. The linear regression implies that only the relative profit to agriculture is identified. Thus, I require a normalization. The forested returns are an outside option with 0 expected return. In the context of the model, this implies no regrowth costs $\phi^F(\omega) = 0$.

Specifying agricultural returns. It remains to parameterize agricultural returns $\pi^A(\omega; s_0)$ in terms of the data. I include two control variables. I denote the lagged share of forest as $f(\omega)$. Lagged dependent variables rationalize low transition rates between land uses. I control for mean travel time to a major city, $c(\omega)$, with coefficient γ_τ.[20] Travel times rationalize low deforestation in forest interiors relative to the frontier. The coefficient ϕ^A on the lagged share of forest $f_{t-1}(\omega)$ indicates the marginal cost of removing additional forest, so that $\phi^A f_{t-1}(\omega)$ represents the empirical switching cost.

An estimating equation. Estimation requires a minimum of two periods of land use data, t and $t-1$. Thus, a naïve pooled OLS regression implied by the landowners' problem, with t given by the year in the data, is:

$$Y_t(\omega) = \gamma \underbrace{[p_{it}^A \eta^A(\omega) + \gamma_\tau c(\omega) + \phi^A f_{t-1}(\omega) + \xi_t(\omega)]}_{\text{empirical relative profit function}} + \lambda_i + \epsilon_t^{FA}(\omega) \tag{11}$$

where potential yields $\eta^A(\omega)$ and travel times $c(\omega)$ are cross-sectional measures.

Wages appear in the landowners' problem, Equation (1), but are fundamentally unobservable. Thus, I introduce a country-level fixed effect λ_i to control for the average level of input prices. I also include other controls for weather (quantiles of heating and cooling degree days), regrowth rates (a cubic polynomial), and, in some specifications, I test richer fixed effects including country-level trends or fixed effects for smaller administrative units (e.g., states or counties).

Two identification challenges. A naïve estimator of Equation (11) implements an ordinary least-squares (OLS) regression. The idealized experiment behind an OLS estimator compares two plots of land which are similar on rich exogenous controls. One has slightly

20. The functional form of travel time is potentially nonlinear. Land close to cities is more expensive to rent due to competition with non-agricultural land use. Land further from cities is less expensive to rent but incurs greater expenses in terms of ferrying goods to market. I thus show results for a simple linear cost, a travel time polynomial, as well as with binned travel times.

better agricultural yield (or a higher agricultural price per ton) than the other. The relative rate of deforestation on the agriculturally better endowed plot of land will determine γ, the responsiveness of deforestation to revenues.

Such estimation of Equation (11) is biased for two reasons. I provide two complementary identification strategies to deal with these threats. First, aggregate differences in unobservable land returns $\xi_t(\omega)$ are correlated with prices and land use shares, generating a simultaneity problem and attenuating estimates of γ. For example, a Brazil-wide supply shock to pastoral land drives agriculturalists to plant soy, raising $\mu^A(\omega)$, *and* lowers prices for soy, lowering p_{it}. Below, I introduce a price instrument.

Second, unobserved shocks to land returns can be serially correlated. Returning to the pastoral shock example, suppose there is a shortage of cattle feeding silos. This is a positive shock to the relative value of cropping through $\xi_{t-1}(\omega)$. Such a shock persists over time, so $\xi_t(\omega)$ is still positive a year later. Such serial correlation will overstate inertia through the lagged dependent variable $f_{t-1}(\omega)$. This lower switching cost is confounded with the depressed pastoral return today $\xi_t(\omega)$. Thus, the positive effect of the shock will be misattributed to lower forest cover $f_{t-1}(\omega)$, resulting in an over-estimate of switching costs. Yet, improperly accounting for these switching costs will bias the estimate of γ. I account for this channel through a spatial first differences exercise.

Instrument for agricultural prices. Resolving simultaneity bias in this supply equation requires a source of exogenous demand variation which only affects agricultural returns through shifts in prices.

I leverage differential exposure to the accession of several countries to the World Trade Organization since 1996. Between 1996-2020, 26 countries access the WTO, gaining favorable terms of trade with existing WTO members. These accessions events provide variation through staggered timing. The exact timing of accession could be subject to some anticipation. Thus, I strengthen this variation further by observing that countries face differential exposure to the same accession shock.

To calculate differential exposure to the common WTO shock, for each country i and accessing partner j, I construct an ex ante measure: the number of potential trade linkages $i \to j$ which are likely to grow in volume after accession. Call this measure V_{ij}.

$$V_{ij} = \sum_{k \in \mathcal{M}_j} \mathbb{1}[X_i^k > 0]$$

for every HS-2 code k demanded by j. I define the set of products $\mathcal{M}_j$ as those imported in non-zero quantities by the accessing country j in the year prior to the first accession, 1995. Using the ITPD-E data, I calculate the fraction of these potentially imported HS-2 level

product codes $k \in \mathcal{M}_j$ which are produced by origin i.

I further limit the set of HS-2 codes to be manufacturing products, as using agricultural products in the instrument could result in invalidity.[21] Then, for each country, I calculate the best-case variety match among the full set of accessing countries, $V_i = \max_j V_{ij}$. If my hypothesized mechanism is correct, a larger exposure measure should imply a larger kink in first-stage agricultural prices post-WTO accession.

The exclusion restriction requires that manufacturing exposure to the accessing country does not shift unobservable agricultural land quality, after controlling for rich observables and fixed effects.[22] For a concrete example of an exclusion violation: more exposed countries are also more intensive miners, which worsens soil quality unobservably. For the instrument to be relevant, accession's effect on land use is mediated by the *relative price* of agriculture and non-agricultural inputs.

Spatial differences. To identify γ in the face of serial correlation in unobservables, I use a spatial first differences strategy (modifying the strategy in Druckenmiller and Hsiang 2019). Define $\{x_\omega, y_\omega\}$ as the coordinates for plot ω in Cartesian space. Define $d\omega$ as the neighboring plot at $\{x_\omega, y_\omega\} + \vec{e}$ for a step $\vec{e}$ in some direction (e.g., in the x-direction, $\vec{e} = \{1, 0\}$). I assume that neighbors share a component of their unobservable return, $\xi(\omega)$. For example, the pastoral suitability or neighborhood amenities of plot ω are correlated with the same of its neighbor, plot $d\omega$. Then, I assume a persistent component of the structural error is correlated with unobservable realizations nearby, up to the AR(1) coefficient ρ:

$$\xi_t(\omega) = \rho \xi_t(d\omega) + v_t(\omega)$$

where the remaining portion of the unobservable, $v_t(\omega)$, is as-good-as-random. This assumption motivates a spatial quasi-difference to remove the effect of the persistent unobservable. Defining $\breve{x}_{it} = \rho x_{it}(\omega) - x_{it}(d\omega)$,

$$\breve{Y}_t(\omega) = \gamma[p_{it}^A \breve{\eta}^A(\omega) + \gamma_\tau \breve{c}(\omega) + \phi^A \breve{f}_{t-1}(\omega)] + \breve{\epsilon}_t^{FA}(\omega) + v_t(\omega) \tag{12}$$

21. Specifically, serial correlation in agricultural supply shocks can result in invalidity. Suppose for example that I use the exposure of Brazilian agriculture to Ecuadorean agriculture as the instrument of choice. If Brazilian agricultural exposure to Ecuador is relatively high due to a contemporaneous unobserved technological innovation in soybean seeds, this innovation would affect the unobservable land quality today. Further, practically, most countries export some volume of agricultural goods, so there is little variation in the variety measure.

22. Borusyak and Hull (2020) argue that for such non-random exposures to exogenous shocks, one ought to control for the expected instrument. I do so by including a fixed effect at the level of the matched "accessing partner" level which captures the average exposure of matched countries.

Spatial quasi-differencing thus removes the persistent unobservable by assumption, leaving only the innovation process v_{it}. Using my previous price instrument, I estimate Equation (12) using a moment condition.

$$\mathbb{E}[(\breve{\epsilon}_t(\omega) + v_t(\omega))V_i] = 0.$$

Intuitively, neighboring units (10 km apart) tend to face similar input prices, so that a spatial first difference effectively cancels out this common component of returns. Observables like $p_i^A \eta^A(\omega)$, meanwhile, have relatively rich local variation remaining to identify γ. More precisely, the identifying assumption underpinning SFD is that the average gradient in agricultural returns is uncorrelated with omitted variables.[23]

Summary. I aim to estimate the dollar cost of land use change, which is pinned down by the parameter γ. I propose two complementary identification strategies in this section. First, a two-stage least squares estimator which remedies simultaneity between prices and output decisions. Second, a spatially differenced estimator, combined with the previous IV, which treats inertia due to serially correlated unobservables.

4.2 Demand parameters and trade barriers

I estimate trade barriers and trade elasticities from data on agricultural and manufacturing trade flows. To estimate trade barriers, I use a ratio estimator in the style of Head and Ries (2001). I then estimate σ^A and σ^M largely following Costinot, Donaldson, and Smith (2016).

4.2.1 Trade barriers

I employ a ratio estimator to back out trade barriers. This estimator captures average barriers across pairs of countries (i, j). I estimate a separate preference-weighted trade barrier in each sector $\bar{T}_{ij}^h = (\theta_{ij}^h T_{ij}^h + \theta_{ji}^h T_{ji}^h)$ through the fixed effects of the following regression:

$$\log\left(\frac{X_{ij}^h X_{ji}^h}{X_{ii}^h X_{jj}^h}\right) = \delta_{ij}^h + \epsilon_{ij}^h \tag{13}$$

where X_{ij}^h, defined in Section 3.2, refers to the trade *flows* of sectoral good h ($h = A$ indicates agricultural goods) between origin i and destination j. Using the Armington demand system,

23. I prefer quasi-differencing to an overt spatial first difference because satellite outcomes are measured with error. Further, as deforestation is a rare event, temporal AR(1) approaches return estimates of ρ close to 1 even over longer panels. The OLS estimator is no longer asymptotically normal at these unit roots due to nonstationarity (e.g., shock effects are increasing in time) (Arellano and Bond 1991).

the fixed effects estimate the average preference-weighted trade barrier between i and j: $\delta_{ij}^h = (1-\sigma^h)\log \bar{T}_{ij}^h$. With a value of the trade elasticities σ^h, the preference-weighted average trade barrier is identified. To account for significant missing or mis-measured agricultural trade data, I set the barrier between (i,j) to $\bar{T}_{ij} = \max\{\bar{T}_{ij}^A, \bar{T}_{ij}^M\}$. My main specification estimates trade barriers off of trade flows from the year 2016. I do not observe trade in the Democratic Republic of the Congo, Romania, and Serbia: I set their incoming trade barrier for exporter i to the 90th percentile of T_{ij} among non-missing observations.

4.2.2 Demand elasticities

To obtain the elasticities σ^A, σ^M, I use a panel gravity approach (Head and Mayer 2014).
Gravity: first stage. Total expenditures in country (destination or importer) j on goods from country (origin, or exporter) i in the manufacturing sector are given by expenditure shares from Equations (A.1) and (A.2) multiplied by income:

$$E_{ij}^h = \lambda_j^{i|h}\lambda_j^h Y_j$$

Taking logs of this expenditure equation and grouping terms, I obtain a structural equation which relates volumes of trade flows to trade barriers:

$$\log E_{ij}^M = (1-\sigma^M)\log\frac{T_{ij}}{\theta_{ij}^M} + \overbrace{(1-\sigma^M)\log\frac{w_i}{\bar{\eta}_i^M}}^{\delta_i^{X,M}} + \overbrace{[(\sigma^M-\sigma)\log P_j^M + \log Y_j + (\sigma-1)\log(\theta_j^M P_j^C)]}^{\delta_j^{M,M}} \quad (14)$$

The next equation groups these terms into three components: a bilateral component, an origin-market i component, and a destination market j-component. Additionally, I introduce a time subscript t, which allows markets (i,j) to observed in the panel:

$$\log E_{ijt}^M = (1-\sigma^M)\log\frac{T_{ij}}{\theta_{ij}^M} + \delta_{it}^{X,M} + \delta_{jt}^{M,M} \quad (15)$$

This equation, the gravity "first-stage," provides the structure for a class of gravity estimators as a fixed-effects decomposition of observed trade flows. I refer to $\delta_{it}^{X,M}, \delta_{jt}^{M,M}$ as the

"propensity" of countries i and j to export and import manufacturing, respectively.

Gravity: second stage. The second stage of gravity estimation follows Redding and Venables (2004) and Duranton, Morrow, and Turner (2014). The exporter fixed effect $\delta_{it}^{X,h}$ is a log-linear function of the endogenous price in agriculture $h = A$ and the endogenous wage in manufacturing $h = M$. For the manufacturing sector,

$$\delta_{it}^{X,M} = (1 - \sigma^M) \log \frac{w_i}{\bar{\eta}_i^M}$$

This equation is problematic: it relies on an endogenous and fundamentally unobservable wage. However, in equilibrium, this wage can be derived from a "market access" term: I provide a full derivation in Appendix E.1. The structural market access term is a function of the fixed effects from Equation (15). Summarizing Appendix Equation (A.6),

$$\delta_{it}^{X,h} = -\frac{(\sigma^h - 1)}{\sigma^h} MA_{it}^{X,h} + \frac{(\sigma^h - 1)}{\sigma^h} \log Q_{it}^h$$

where h indicates a sector (in my case, agriculture or manufacturing) and Q_{it}^h is the aggregate supply of good h: Q^A in the agricultural sector and $Q^M = \bar{\eta}^M M_i$ in manufacturing. Converting this to a regression equation,

$$\delta_{it}^{X,h} = -\frac{(\sigma^h - 1)}{\sigma^h} MA_{it}^{X,h} + \beta^h \log X_{it}^h + v_{it}^{X,h} \tag{16}$$

where X_{it}^h are observable inputs into aggregate production. The error term $v_{it}^{X,h}$ represents the unobservable component of production: in agricultural industries, this is land quality ξ_{it}, whereas in manufacturing this is total factor productivity $\log \bar{\eta}_i^M$. In X_{it}^h, for the agriculture-specific regression, I control for lagged agricultural employment, fertilizer usage, and pesticides using data from the FAO. In non-agricultural industries, I control for lags of employment. I further test the model in first differences to limit bias from unobserved inputs to production. As long as input demand is persistent over time, first differencing reduces bias (though, with generated regressors in particular, measurement error could be an issue).

When tracing out the resulting aggregate demand curve, a key challenge comes from unobservable productivity terms correlated with export volumes. High-productivity regions are both higher-priced (higher market access) and export more (on the left hand side). This simultaneity generates attenuation in the coefficient of interest, biasing estimates towards 0. To break this simultaneity issue, I instrument for current market access using lagged market access. This approach relies on a timing assumption, where prices are set after

observing transient demand shocks (Anderson and Hsiao 1981; Scott 2013; Araujo, Costa, and Sant'Anna 2020).

However, just as today's market access is correlated with contemporaneous productivity, lagged market access is correlated with lagged productivity. Then, if there is any serial correlation in productivities, lagged market access is correlated with current productivity. To account for serially correlated productivity, I use a leave-one-out lagged market access instrument which instruments for market access in i with an average market access outside of i. Thus, the instrumental variables exclusion restriction imposes:

$$\mathbb{E}\left[\frac{1}{N-1}\sum_{j\neq i, j\in\{1,\ldots,N\}} MA^h_{j,t-1} v^{X,h}_{it} \,\middle|\, \log Q^h_{it}\right] = 0.$$

This IV requires that supply shocks abroad in the previous period do not drive domestic shocks today. An example of a violation would be a large increase to oil reserves in the US, a large oil producer, which is both serially correlated (oil reserves are higher next year as well) and would drive supply decisions today.

Outer nest parameters. The outer nest elasticity σ governs substitution between manufacturing and agricultural consumption in response to a relative price change. I estimate the model for two values of σ: the Cobb-Douglas case, $\sigma = 1$, and the value of σ in Comin, Lashkari, and Mestieri (2021), $\sigma = 0.5$. I set outer nest parameters θ^h_i to match sectoral expenditure shares to the World Bank's International Comparison Program data. These aggregate quality differences capture countries with an unexplained high market share after controlling for prices (Khandelwal 2010). I discuss a reduced-form nonhomothetic preference in Appendix Section E.3.

5 Empirical Results

5.1 Land use regression

Linear approach. I estimate Equation (11) using OLS and two-stage least squares estimators in Table 2. The resulting land use elasticity has a positive sign, and cost shifters move profits in the expected negative direction. Introducing the variety shifter instrument further increases the magnitude of the estimate. In Appendix D.2, I demonstrate a strong first stage of the variety shifter instrument with an F-statistic of 20.088 and provide evidence in favor of the theorized mechanism.

A value of $\gamma = 1.4$ implies a *short-run* deforestation elasticity of 0.09. However, long-run elasticities (as landowners are myopic) are in line with prior estimates: I return to

this in my discussion of model results. Globally, this value of γ implies a 1% increase in agricultural share requires, on average, \$41.73 per hectare in additional potential revenues. For comparison, Souza-Rodrigues (2019) estimates a figure of \$42/ha in Brazil.

Introduction of a spatial difference. Recall that the spatial differencing implemented in Equation (12) theoretically mitigates confounding between serially correlated, persistent unobservables and the lagged dependent variable. It refines attention to deforestation variation which cannot be explained by ones' neighbors. Back to Table 2, Columns (3) and (4) show spatial quasi-differences in the X-direction without and with instruments, respectively. Spatial differencing does not alter estimated switching cost magnitudes. Further, spatial differences return a fairly similar estimate of the land use elasticity γ which is stable across differencing directions (see Appendix Table A.6). Travel time elasticities are not well-identified by the SFD: travel time does not have meaningful variation between neighbors 10 kilometers apart.

Given aligned estimates of the switching cost in Column (4), the switching cost of deforestation is estimated at \$163 per hectare of forest. For comparison, in Brazil, Araujo, Costa, and Sant'Anna (2020) estimate average switching costs of \$564/ha. Relative to my estimate, the target \$190-per-ton CO_2 tax increases the average cost of deforestation by 91%.

Robustness checks. Results are consistent across alternate yield measures in Appendix Table A.10. Appendix D.4 discusses an alternate instrument and exclusion restriction which delivers a similar coefficient estimate.

One concern is that the unobserved land quality measure $\xi(\omega)$ would be correlated with specification error in the production function. I correlate these land qualities with a number of proxies for input intensity, including population density (Sayre 2023), population growth, distances to ports and measures of nearby port infrastructure (Farrokhi and Pellegrina 2023). I reject statistically significant relationships in aggregate. Quality more strongly correlates with elevation and aridity. Appendix Figure A.13 maps quality in Brazil.

Next, global average parameters can mask important underlying heterogeneity, which in turn can affect predictions regarding counterfactual policies. For example, if the deforestation frontier is systematically less elastic to price interventions like a carbon tax, then a carbon tax will lead to systematically fewer emissions reductions in these regions. An average parameter would miss this heterogeneity.

Heterogeneity results are reported in Appendix Table A.12. Accounting for cross-country heterogeneity in travel or switching costs has a small effect on the coefficient of interest, γ, in columns (1) and (2). I decompose switching cost heterogeneity in Appendix D.7.

In Column (3), I allow for an explicit interaction between revenues and travel times. Results reveal that distant regions are less elastic than regions close to market, as e.g. labor

availability is higher near markets (Foster and Rosenzweig 2021). Moving 1 hour further from market results in a 10% decrease in the elasticity of deforestation.

Scrap and option values of deforestation differ across space, and they may be arbitrarily correlated with returns. I measure secondary forest regrowth, a part of the option value of deforestation. I also indirectly measure the scrap (timber) value of forest through its biomass value. I allow for country-level heterogeneity in the responsiveness to option and scrap values of forest to reflect differences in timber market infrastructure. Such flexible specifications in columns (4) and (5) have small effects on the price coefficient.

Appendix D.8 tests for unobservable heterogeneity in both preference parameters and switching costs. In a key insight from this analysis, countries with more smallholder farming are less elastic to agricultural returns (γ is smaller). Smallholders face larger outside shocks ϵ and thus attend less to revenues.

A reduced form version of the full model experiment. To further contextualize the magnitude of γ, I calculate a revealed preference measure of agricultural returns on land deforested in 1982-2016:

$$\Delta\text{Returns, USD}_{1986-2016}(\omega) = \frac{1}{\hat{\gamma}}\Delta \overbrace{\left[\log\left(\frac{\mu^A(\omega)}{\mu^F(\omega)}\right)\right]}^{\text{Data on land use shares}} := \underbrace{\Delta[\pi^A(\omega; s_0(\omega)) - \pi^F(\omega; s_0(\omega))]}_{\text{structural interpretation}}$$

I compare this measure with the environmental deadweight loss at a cost of carbon of \$190 as $190 \times d(\omega)$. If a plot is reforested, I calculate its 30-year biomass gain consistent with measured regrowth rates $r(\omega)$ in the data, e.g., $t^{SCC} \times 30 \times r(\omega)$.

I plot agricultural returns and deforestation emissions costs in Appendix Figure A.14. Land above the red 45 degree line had larger environmental costs than agricultural gains. From this scatterplot alone, a data-driven estimate suggests 70% of land was deforested inefficiently. In Section 2.2, 56% of land had higher emissions than yields in *quantities*. Then, high-yield land received lower prices.

More precisely, the average agricultural return on deforested land since 1982 was \$36.14 per hectare, while the average emissions cost of this deforestation \$161.12 per hectare (at a \$190 social cost of carbon). The average agricultural return is 22% of the emissions cost.[24] For comparison, using country-level administrative data, global average value added per hectare is \$58/ha in 2019.[25] Deforested land is negatively selected on agricultural productivity.

24. For context, the average value of γ which would set agricultural returns on average equal to emissions costs is 0.33, an elasticity of 0.02. This elasticity is rejected by all specifications.

25. I calculate global average value-added per hectare using a combination of FAO and World Bank data

Relying on this scatterplot alone has two drawbacks. First, simply dropping this 70% of deforestation is a large intervention. This large intervention would change prices and wages, and change agricultural returns, rendering 70% a seeming overestimate. Second, this 70% of deforestation includes arbitrary land quality distortions $\xi(\omega)$ which have unclear sources: to claim this 70% figure is correct, we must also make the unsavory assumption that these agricultural land distortions are fully "efficient". Then, 70% can be an understatement.

5.2 Demand estimation

In Appendix Section E.2, I report a summary of the first-stage distance elasticities of trade flows against bilateral distance. The first-stage distance elasticities are consistent with a long literature on the effect of distance on trade flows (Head and Mayer 2014; Redding and Venables 2004), and replicate the differences in elasticities of trade measured in weight and value reported in Duranton, Morrow, and Turner (2014). In my discussion here, I focus on the second-step estimators for trade elasticities σ^A, σ^M. Results are reported in Tables 3 (agriculture) and 4 (manufacturing).

The ordinary least squares estimator for the agricultural second stage is reported in column (1) of Table 3. The implied elasticity of substitution σ^A of 9 is in the range of values seen in the literature, between 4-11 (Head and Mayer 2014). This estimate is robust to first differences in columns (2-4) and the lagged market access instrument in column (3).

The manufacturing second stage in Table 4 follows a similar pattern. I have fewer observable controls for manufacturing production. Accordingly, first differencing plays a larger role in correct identification. Interpreting this channel, countries with higher prices also have higher unobserved factor endowments, meaning that they export more (e.g., aside from labor, high-price countries have better infrastructure). Absent a first difference, these unobservables attenuate σ^M. This first differencing attenuation is larger in magnitude than the attenuation corrected by the lagged market access IV. Then, the persistent-across-time component of unobserved TFP $\bar{\eta}_i^M$ is much larger in magnitude than the idiosyncratic component. Then, the instrument has little remaining variation to partial out.

Both results are robust to a variety of specifications. Dropping years prior to the dissolution of the Soviet bloc has no impact on estimates. Adding heterogeneous time trends in within-country travel times, agricultural productivity, and lagged 1980 biomass has no effect.

---

as:

$$\text{Value added per ha} = \frac{\sum_{i=1}^{J} \text{value add per worker}_i \times \text{total employment}_i \times \text{share employed in agriculture}_i}{\sum_{i=1}^{J} \text{hectares land in agriculture}_i}$$

Alternate instruments also do not change the coefficient meaningfully. Finally, dropping individual countries (of 155 in the sample, an approximate influence perturbation exercise) results in a lower bound on the agricultural elasticity of 8.9 (0.8) and a lower bound on manufacturing elasticity of 5.4 (0.4).

6 Model results

6.1 Model fit and status quo deforestation

I calibrate the model on land use maps from 1982 to 2000 and test its calibration on land use data from 2000 to 2016. Two parameter vectors, each $J \times 1$ (two parameters per country), each target a $J \times 1$ dimensional moment. Preference shifters for agriculture θ_i^A set sectoral expenditure shares in Equation (A.2) to their values in the World Bank ICP data for 2010. Agricultural labor shifters a_i in Equation (1) are set such that the share of labor in agricultural value added matches FAO data in 2010. Details are provided in Appendix B.1.

With the specified calibration, Appendix Table A.23 shows that the model captures the correlation between damages and observables: agricultural productivity $\eta^A(\omega)$ and transportation costs. Capturing these correlations helps to pin carbon tax pass-through through the plot ω-level elasticity $s(\omega)$ in Equation (10). Using only information up to the year 2000, the model has an overall R^2 of 0.66 for plot-level emissions from 2000-2016. One interpretation of this correlation is that the remaining 34% of variation in emissions comes from other plot-level amenities ($\xi(\omega)$ in my empirics). Appendix Figure A.11 summarizes the match across biomes, demonstrating the model suits forested biomes particularly well.

Baseline estimates also match the distribution of agricultural production across countries. Plotted in Appendix Figure A.12, the model achieves an R^2 of 0.58 with cross-country production data from the FAO (both crop quantities are measured in total tonnage).

6.2 Counterfactual 1: a global Pigouvian tax

In this section, I quantify how much deforestation from the "business-as-usual" scenario from 1982-2016 above ought to be avoided. I organize results into global and disaggregated (e.g., country and plot-level) production, prices, and welfare.

Effects on global quantities of food and emissions. The results of the Pigouvian tax counterfactual for a range of parameter assumptions are enumerated in Table 5. For my preferred counterfactual specification, I employ an outer nest elasticity of $\sigma = 0.5$. In this preferred specification, emissions from deforestation fall by 97%. This decrease in deforestation corresponds to a 4% increase in total carbon sequestered by forest basins. Even

after accounting for trade costs and general equilibrium interactions with the manufacturing industry, the deforestation distortion is large enough that all but a few regions of the world were too expensive to deforest at a $190 SCC, totaling 88 Gigatons of inefficient carbon between 1982-2016. Inefficient emissions exceed global carbon emissions in 2022. Across specifications, at least 95% of business-as-usual emissions are abated under a $190 tax. Appendix Table A.25 shows that the precise choice of yield aggregation does not affect this conclusion: using calorie-weighted, high-input yields as in Costinot, Donaldson, and Smith (2016) results in similar magnitudes.

Reallocating land through the tax converts 13% of business-as-usual agricultural land area to forested land area. This reallocation results in 7% $(6, 9)$ lower total global agricultural yields. The land area lost tracks almost directly with descriptive Figure 2, where 10% of global land area generated almost the entirety of global deforestation emissions. While in the data, this land generates around 13% of food production, the model projects an equilibrium reduction in food production of only 7%. Then, 6% of food production is substituted by production elsewhere, in part through international trade.[26]

Plot-level results: production. Next focusing on the distribution of food production, Figures 3a and 3b map food production and emissions under the Pigouvian tax scenario. Several forests are worth deforesting, particularly: the Atlantic Forest in Brazil (not the part of the Amazon which was involved in high-profile burning since 2018, but a forest involved in an earlier surge of deforestation in Brazil in the 1980), Mexican forests, East African forests, and the non-Indonesian southeast Asian forest basin. These regions have sufficiently high agricultural production to justify a relatively low emissions cost. Average yields on plots ω which retain at least 10% of business-as-usual emissions are 16 tons per hectare, whereas average starting biomass on this land is 6 tons per hectare. In contrast, plots which emit less than 10% of business-as-usual CO_2 yield 11.6 tons per hectare on average, but emit 30 tons of CO_2 per hectare.

Effects on prices. I now parallel the dramatic result in quantities with one in prices. Despite a large reduction in emissions from deforestation, food prices do not skyrocket. The average pass-through rate is 2.17% $(2.1, 2.3)$ onto prices, and 0.20% $(0.18, 0.23)$ for wages.[27]

Low pass-through results from an elastic agricultural demand ($\sigma^A = 9.1$, or an elasticity

26. The net loss of food production implies that consumers are substituting away from food. There are several sources of food which are substitutable: cattle feeds, corn and soy-derived inputs to manufacturing, and seed oils are all common downstream uses of agricultural products involved in deforestation.

27. Aggregate global pass-through rates are derived by taking a GDP-weighted sum of prices relative to business-as-usual: $\bar{p} = \sum_{i=1}^{J} W_i^{BAU}(p_i^{TAX} - p_i^{BAU}) / \sum_{i=1}^{J} W_i^{BAU}$. Thus, this number can be interpreted as the percent increase in agricultural prices faced on the average dollar of real income.

of 0.89) and an inelastic supply of agricultural land (0.1).[28] Without accounting for wage pass-through and the spatial correction in Equation (10), the classic pass-through equation from Jenkin (1872) yields 11% pass-through. Thus, in partial equilibrium landowners pay \$0.90 per dollar of taxes. In Table 5, wages account for a very small fraction of the gap between this partial equilibrium pass-through and actual GE pass-through of 2.17%.

The remaining pass-through gap comes from plot-level heterogeneity. High-emissions firms bearing the cost of the tax are not the most productive firms, yielding to a lower spatial correction in Equation (10). I verify this story by plotting the spatial correction term against pass-through rates in Appendix Figure A.1. Based on the R^2 from the line of best fit in this figure, cross-country variation in subnational distributions of agricultural yields and emissions explain 72% of variance in pass-through across countries. Remaining variation comes from tariffs on the demand side and differences in levels of productivity on the supply side, which affect elasticities.

Robustness checks for other values of (σ, σ^A) in Table 5 support these conclusions. Cobb-Douglas preferences $\sigma = 1$ lower pass-through relative to the preferred gross complements case $\sigma = 0.5$ as consumers more readily substitute away from food and towards manufacturing in response to relative price changes. A less elastic agricultural trade elasticity of $\sigma^A = 4$ relative to my estimate of $\sigma^A = 1$ raises pass-through.

Effects on global welfare. I summarize welfare consequences as tax-induced changes in real expenditure per ton of avoided emissions. These average abatement costs, defined formally in Equation (9), can be used to benchmark deforestation relative to other large climate policy interventions. Table 5 finds that implementing \$190 Pigouvian tax costs 4% of business-as-usual GDP. Converting this to equivalent variation per ton, average global abatement costs are \$106 (100, 122) per ton.

Distributional effects on welfare. The tax does not change the cross-country distribution of agricultural wealth. Landowners in wealthier countries lose on average a similar percentage of income to landowners in poorer countries at the \$190 tax rate (see Appendix Figure A.16).

However, the tax is regressive within countries. High-emissions landowners also have low returns, meaning that they are poorer producers. In Appendix Figure A.17, the gap between the richest and poorest quartiles of the distribution of farmers within the average country is widening. Dispersion rises by nearly 9% relative to its starting value.

28. In Appendix B.4, I compare the model-implied crop acreage-to-crop price elasticity to values from the literature, finding that they are broadly consistent.

6.3 Decomposing mechanisms in counterfactual 1

My modeling approach highlights three channels: international trade costs, comparative advantage in non-agriculture, and plot-level heterogeneity in costs. I re-run the \$190/ton global Pigouvian tax counterfactual but successively shut down each channel in turn. I describe each decomposition and its aggregate welfare impacts in Appendix Table A.2. I quantify how each channel affects remaining food (6% reduction at baseline) and emissions (97% reduction at baseline) under the Pigouvian tax in Figure 5.

Trade. In welfare terms, from Appendix Table A.2, free trade reduces the deadweight loss of the tax by 40%, nearly doubling the social surplus from Pigouvian taxes. In Figure 5, status quo trade barriers account for half of the remaining emissions in the Pigouvian outcome and 30% of Pigouvian food production losses. Free trade enables easier reallocation across borders and demand acts as-if more elastic. Thus, under the Pigouvian tax with free trade, agricultural production readily shifts towards highly productive farmland in the US, Canada, and the EU. Consumers in tropical countries gain from lower production costs and higher yields abroad. Thus, trade plays a key role in environmental policy effectiveness (Le Moigne et al. 2024).

Autarky decreases the deadweight loss of the Pigouvian tax for two reasons. First, in autarky, large relative deadweight losses in the tropics do not affect prices in the US or the EU. These large low-biomass economies thus face limited welfare effects from trade. Second, agricultural production is not reallocated across borders, keeping prices high in the tropics. Inframarginal landowners in the tropics capture larger windfalls. Even within the tropics, losses to producers are lower (though consumers pay higher food prices).

Cross-sector spillovers and wages. This experiment shuts down tax-induced equilibrium wage adjustments. From Appendix Table A.2, at least 36% of deadweight loss on landowners comes from rising nominal wages and associated rising real manufacturing prices. Exogenous wages extensify food production in the US and China, where non-agricultural industries have a notable comparative advantage.

Production costs. I set production costs $c(\omega)$ across plots ω to the tenth percentile of the within-country distribution in the final decomposition exercise. Agricultural production becomes cheaper for 90% of the world. This choice is arbitrary: for different choices of the distribution of production costs, deadweight loss can be higher or lower. This channel has the smallest net effect on welfare and production.

General equilibrium. Finally, in Figure 5, I consider the importance of a general equilibrium approach, writ large. A first-order approximation of the effects of a Pigouvian tax

would simulate land use shares from the landowners' discrete choice problem without considering price changes. This first-order approach overstates emissions reductions relative to the full model and also overstates agricultural costs per ton. Tax-induced profit losses lead to shrinking agricultural supply, but excess demand-driven rising agricultural prices no longer compensate some landowners. Numerically, endogenous prices explain 50% of remaining emissions from deforestation.

6.4 The global average marginal abatement cost curve.

In this section, I explore aggregate welfare costs as a function of the carbon tax rate. I quantify the costs of deforestation using the *average* global marginal abatement cost, or the $AMAC$ curve, in response to a tax rate. Globally, abatement costs are equal to equivalent variation per ton of emissions avoided. Thus, marginal abatement costs take the (numerical) derivative of this abatement cost curve.[29]

Figure 4 illustrates the global $AMAC$ curve for the entire economy. Average marginal abatement costs are generally higher than the landowners' private cost of cutting emissions (given by the tax rate). These larger aggregate costs reflect a producer channel and a consumer channel. On the producer side, rising agricultural prices raise agricultural returns. On the consumer side, consumers substitute agricultural consumption from their preferred basket to new, deforestation price-adjusted baskets. I illustrate the contribution of each channel in Appendix Figure A.18.

The cost curve reflects the changing "technology" involved in reducing emissions from deforestation. Initially, carbon pricing reallocates low-returns agriculture. These low-hanging fruit produce little aggregate yield, which implies limited welfare costs from reallocation. Up to a tax rate of \$8, Brazil is the largest supplier of emissions reductions. However, at \$8, Russian supply overtakes the Brazilian supply of emissions, inducing a kink in the abatement cost curve. Russian deforestation has higher returns per ton as Boreal agricultural yields are fairly concentrated in limited land area. Finally, at a tax rate of \$70, China becomes the largest supplier of emissions reductions, but these reductions are extremely costly.

6.5 Counterfactual 2: Incomplete regulations

In this counterfactual, I set country-level tax rates to their highest recorded levels across taxes and Emissions Trading System prices recorded in the Carbon Pricing Dashboard of the World Bank. I then apply these incomplete and spatially heterogeneous carbon prices to

29. AMAC is distinct from the marginal abatement cost of the tax, which for any landowner ω with non-zero emissions is set to match the tax rate t.

deforestation. The average regulated plot of land faces a tax of $18.38. Partial regulation, in addition to mechanically covering less area, can generate spillovers from regulated to less-regulated or unregulated regions (leakage, henceforth). Further, if policies are not targeted to low-returns deforestation, emissions reductions can be cost ineffective.

The results are stark: current tax rates abate only 5% of business-as-usual emissions (relative to 97% from a global uniform tax). Emissions fall in the *wrong* places relative to the Pigouvian global standard, as plotted in Figure 6. For example, the EU stays close to its Pigouvian levels of abatement, but it has low status quo emissions.

Deforestation can spill over borders absent international policy coordination. The aggregate leakage rate is 11%: for every 100 tons of emissions reductions in regulated countries, deforestation-related emissions in unregulated countries increase by 11 tons. The US and Canada deforest 6.7% and 8% more, respectively, as their agricultural production is the most direct substitute to the EU. Brazil emits 3.8% more CO_2 from deforestation, a more modest rate but on a much higher base. Turning to welfare, this patchwork of domestic deforestation taxes increases global real expenditures (in equivalent variation terms, $28 billion USD).[30] The primary beneficiaries from this policy are the US and Canada, both of whom profit off of excess agricultural demand in the EU.

Leakage can benefit countries who regulate less (e.g., set a lower carbon tax). To understand the role of this internal leakage, I solve for a single tax rate across all countries with a non-zero carbon price. The single price is set to achieve the same aggregate emissions reductions as the baseline partial regulation scenario. From the average plot-level tax rate of $18, this cost-effective policy brings costs down to $5.86 per ton on the margin. Thus, in higher-tax countries (e.g., in the EU) the marginal prevented deforestation had higher agricultural returns than the marginal deforestation which actually takes place in other regulating countries (e.g., in a lower tax regime like Mexico).

Finally, I compare the policy to a global cost effective policy which achieves an equivalent reduction in global emissions. The cost effective policy is a global Pigouvian tax, and would cost only $0.19 per ton. Compared with the free-riding corrected cost of the policy of $5.86 per ton, this result suggests that we could obtain the same emissions reductions by retargeting across unregulated countries at 3% of the cost. Similar to the prior quantification, this exercise suggests that unregulated regions have most of the "low-hanging fruit," so establishing broader carbon markets is highly beneficial.

30. Welfare can increase under partial regulations in an economy where there are distortions from trade barriers (Bai, Jin, and Lu 2019). Here, welfare increases because non-taxing countries benefit from agricultural production shifting towards productive land in the US.

6.6 On co-benefits from carbon taxes

In this section, I construct a "co-benefit" curve to marginal (potential) species habitat protection using a measure of rarity-weighted species richness. I define the co-benefit curve as the change in emissions per hectare of habitat of rare species. I briefly describe how I obtain this curve from the model.

I use an implicit price t^b on that habitat to back out the full curve, similar to the global AMAC curve exercise in Section 6.4. For a measure of the importance of habitat on plot ω $b(\omega)$, landowners will face a modified agricultural return as:

$$\pi_\epsilon^A(\omega; s_0(\omega), t^{SCC}, t^b) = \pi_\epsilon^A(\omega; s_0(\omega), t^{SCC}, 0) - t^b b(\omega)$$

Then, the co-benefit curve is a ratio between the habitat value of lost agricultural land and its emissions value:

$$\text{Co-benefits}(t; t^b) = \frac{\sum_{\omega=1}^N d(\omega)[\mu^A(\omega; s_0(\omega), t, t^b) - \mu^A(\omega; s_0(\omega), t, 0)]}{\sum_{\omega=1}^N b(\omega)[\mu^A(\omega; s_0(\omega), t, t^b) - \mu^A(\omega; s_0(\omega), t, 0)]}$$

The co-benefit curve is largest when the emissions on high-habitat land is also large. It formalizes a simple intuition, which is that the co-benefits from reducing cabon emissions on biodiversity depends on how much biodiversity spatially coincides with carbon value. I remain agnostic about the true social shadow value t^b which maximizes welfare.

Figure A.21 shows my co-benefit measure for a range of 25 shadow values t^b from \$0.001 to \$1000 per species count. Each panel of the figure uses a different measure of habitat value $b(\omega)$ which emphasizes a different set of threatened species from the IUCN Red List. For small shadow values of habitat loss, Amphibian habitat costs the most in terms of carbon tonnage. This cost comes from a geographic difference: amphibian hotspots in the data are concentrated in Central America – notably Panama – relative to dense biomass in South America. Thus, there is an increased "substitution" between policies which target only biomass and policies which target protection of amphibian habitats.

When focusing on the rarest species, Southeast Asian forests (primarily Indonesian, Vietnamese, and and Malaysian forests) are differentially less protected. Their deforestation is environmentally most costly from a biomass perspective. In contrast, Brazilian deforestation is unaffected by additionally valuing rare species habitats, suggesting that co-benefits in the Amazon are quite high from carbon protection alone.

7 Discussion

There are three key takeaways from this work. First, deforestation-related carbon emissions can be overwhelmingly reduced with limited food costs. This overwhelming reduction in deforestation is an extremely robust result: it accounts for trade policy, agricultural policy, inputs costs, and endogenous labor costs. Much of historical deforestation was thus "low-hanging fruit" with low agricultural returns and high emissions. Deforestation on low-hanging fruit also implies that global anti-deforestation policy is inequality-increasing.

Second, the geography of deforestation depends on distortions from trade, differences in non-agricultural productivity, and within-country cost variation. All three forces significantly increase the cost of undertaking global deforestation policy by reallocating agricultural land away from regions with a measurable comparative advantage.

Finally, current carbon pricing needs to coordinate any extensions towards agriculture. Current, uncoordinated carbon pricing interventions would save a mere 5% of global deforestation by targeting high-value agriculture, resulting in high costs. Withi political constraints on carbon pricing interventions, and lackluster track records in offsets markets, this 5% emissions reduction can still be achieved at lower costs by coordinating carbon taxes.

Indeed, global deforestation mitigation strategies must take into account the joint nature of the deforestation decision: not deforesting implies a loss of agricultural production for local communities. Otherwise, high-value – often subsistence, in this context – agriculture is lost to conservation efforts. I highlight the fundamental tension between food production and deforestation on a global scale.

References

Abman, Ryan, and Clark Lundberg. 2020. "Does Free Trade Increase Deforestation? The Effects of Regional Trade Agreements." Publisher: The University of Chicago Press, *Journal of the Association of Environmental and Resource Economists* 7, no. 1 (January): 35–72. ISSN: 2333-5955, accessed February 22, 2024. https://doi.org/10.1086/705787. https://doi.org/10.1086/705787.

Adamopoulos, Tasso, and Diego Restuccia. 2022. "Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data." *The Review of Economic Studies* 89, no. 4 (July): 1629–1653. ISSN: 0034-6527, accessed April 10, 2023. https://doi.org/10.1093/restud/rdab059. https://doi.org/10.1093/restud/rdab059.

Anderson, T. W., and Cheng Hsiao. 1981. "Estimation of Dynamic Models with Error Components." Publisher: [American Statistical Association, Taylor & Francis, Ltd.] *Journal of the American Statistical Association* 76 (375): 598–606. ISSN: 0162-1459, accessed July 14, 2022. https://doi.org/10.2307/2287517. https://www.jstor.org/stable/2287517.

Araujo, Rafael, Francisco J. M. Costa, and Marcelo Sant'Anna. 2020. *Efficient Forestation in the Brazilian Amazon: Evidence from a Dynamic Model* [in en]. Technical report. Publication Title: SocArXiv. December. Accessed June 16, 2022. https://ideas.repec.org/p/osf/socarx/8yfr7.html.

Arellano, Manuel, and Stephen Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." Publisher: Wiley-Blackwell, *The Review of Economic Studies* 58 (2): 277–297.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" [In en]. *American Economic Review* 102, no. 1 (February): 94–130. ISSN: 0002-8282, accessed August 25, 2022. https://doi.org/10.1257/aer.102.1.94. https://www.aeaweb.org/articles?id=10.1257/aer.102.1.94.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The Geography of Trade and Technology Shocks in the United States." *American Economic Review* 103, no. 3 (May): 220–25. https://doi.org/10.1257/aer.103.3.220. https://www.aeaweb.org/articles?id=10.1257/aer.103.3.220.

Bai, Yan, Keyu Jin, and Dan Lu. 2019. *Misallocation under trade liberalization.* Technical report. National Bureau of Economic Research.

Balboni, Clare, Aaron Berman, Robin Burgess, and Benjamin A. Olken. 2023. "The Economics of Tropical Deforestation." Publisher: Annual Reviews, *Annual Review of Economics* 15, no. 1 (September): 723–754. ISSN: 1941-1383, accessed March 7, 2024. https://doi.org/10.1146/annurev-economics-090622-024705. https://doi.org/10.1146/annurev-economics-090622-024705.

Bergquist, Lauren F, Benjamin Faber, Thibault Fally, Matthias Hoelzlein, Edward Miguel, and Andrés Rodríguez-Clare. 2022. *Scaling Agricultural Policy Interventions.* Working Paper 30704. Series: Working Paper Series. National Bureau of Economic Research, December. https://doi.org/10.3386/w30704. http://www.nber.org/papers/w30704.

Borchert, Ingo, Mario Larch, Serge Shikher, and Yoto V. Yotov. 2021. "The International Trade and Production Database for Estimation (ITPD-E)." *International Economics* 166:140–166. ISSN: 2110-7017. https://doi.org/https://doi.org/10.1016/j.inteco.2020.08.001. https://www.sciencedirect.com/science/article/pii/S2110701720302602.

———. 2022. *The international trade and production database for estimation - release 2 (ITPD-E R02),* July. https://sussex.figshare.com/articles/report/The_international_trade_and_production_database_for_estimation_-_release_2_ITPD-E_R02_/23496356.

Borusyak, Kirill, and Peter Hull. 2020. *Non-random exposure to exogenous shocks: Theory and applications.* Technical report. National Bureau of Economic Research.

Borusyak, Kirill, and Xavier Jaravel. 2024. "Are trade wars class wars? The importance of trade-induced horizontal inequality." *Journal of International Economics* 150:103935. ISSN: 0022-1996. https://doi.org/https://doi.org/10.1016/j.jinteco.2024.103935. https://www.sciencedirect.com/science/article/pii/S002219962400062X.

Cain, Lucas, Danae Hernandez-Cortes, Christopher Timmins, and Paige Weber. 2024. "Recent Findings and Methodologies in Economics Research in Environmental Justice." _Eprint: https://doi.org/10.1086/728100, *Review of Environmental Economics and Policy* 18 (1): 116–142. https://doi.org/10.1086/728100. https://doi.org/10.1086/728100.

Carleton, Tamma, Levi Crews, and Ishan Nath. 2023. *Agriculture, Trade, and the Spatial Efficiency of Global Water Use.* Working Paper.

Christensen, Peter, and Christopher Timmins. 2022. "Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice." _Eprint: https://doi.org/10.1086/720140, *Journal of Political Economy* 130 (8): 2110–2163. https://doi.org/10.1086/720140. https://doi.org/10.1086/720140.

Clausing, Kimberly A., and Catherine Wolfram. 2023. "Carbon Border Adjustments, Climate Clubs, and Subsidy Races When Climate Policies Vary." *Journal of Economic Perspectives* 37, no. 3 (September): 137–62. https://doi.org/10.1257/jep.37.3.137. https://www.aeaweb.org/articles?id=10.1257/jep.37.3.137.

Comin, Diego, Danial Lashkari, and Martí Mestieri. 2021. "Structural change with long-run income and price effects." Publisher: Wiley Online Library, *Econometrica* 89 (1): 311–374.

Costinot, Arnaud, Dave Donaldson, and Cory Smith. 2016. "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields around the World." Publisher: The University of Chicago Press, *Journal of Political Economy* 124, no. 1 (February): 205–248. ISSN: 0022-3808, accessed April 10, 2023. https://doi.org/10.1086/684719. https://www.journals.uchicago.edu/doi/abs/10.1086/684719.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings." *American Economic Review* 105, no. 2 (February): 678–709. https://doi.org/10.1257/aer.20121656. https://www.aeaweb.org/articles?id=10.1257/aer.20121656.

Currie, Janet, John Voorheis, and Reed Walker. 2023. "What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality." *American Economic Review* 113, no. 1 (January): 71–97. https://doi.org/10.1257/aer.20191957. https://www.aeaweb.org/articles?id=10.1257/aer.20191957.

Curtis, Philip G., Christy M. Slay, Nancy L. Harris, Alexandra Tyukavina, and Matthew C. Hansen. 2018. "Classifying drivers of global forest loss." Publisher: American Association for the Advancement of Science Section: Report, *Science* 361, no. 6407 (September): 1108–1111. ISSN: 0036-8075, 1095-9203, accessed May 26, 2021. https://doi.org/10.1126/science.aau3445. https://science.sciencemag.org/content/361/6407/1108.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2007. "Unbalanced trade." Publisher: American Economic Association, *American Economic Review* 97 (2): 351–355.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review* 109, no. 12 (December): 4178–4219. https://doi.org/10.1257/aer.20180279. https://www.aeaweb.org/articles?id=10.1257/aer.20180279.

Dominguez-Iino, Tomas. 2023. *Efficiency and Redistribution in Environmental Policy: An Equilibrium Analysis of Agricultural Supply Chains.* Working Paper, May. Accessed July 19, 2023. https://doi.org/https://www.tomasdomingueziino.com/.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2018. "Measuring the Bias of Technological Change." Publisher: The University of Chicago Press, *Journal of Political Economy* 126, no. 3 (June): 1027–1084. ISSN: 0022-3808, accessed October 30, 2024. https://doi.org/10.1086/697204. https://doi.org/10.1086/697204.

Druckenmiller, Hannah, and Solomon Hsiang. 2019. *Accounting for Unobservable Heterogeneity in Cross-Section Using Spatial First Differences.* Working Paper 25177. Series: Working Paper Series. National Bureau of Economic Research, June. https://doi.org/10.3386/w31008. https://www.nber.org/system/files/working_papers/w25177/w25177.pdf.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner. 2014. "Roads and Trade: Evidence from the US." Publisher: Oxford University Press, *Review of Economic Studies* 81 (2): 681–724.

Efron, Bradley. 1987. "Better bootstrap confidence intervals." Publisher: Taylor & Francis, *Journal of the American Statistical Association* 82 (397): 171–185.

Farrokhi, Farid, Eliot Kang, Heitor S. Pellegrina, and Sebastian Sotelo. 2023. *Deforestation: A Global and Dynamic Perspective.* Working Paper.

Farrokhi, Farid, and Heitor S. Pellegrina. 2023. "Trade, Technology, and Agricultural Productivity." _Eprint: https://doi.org/10.1086/724319, *Journal of Political Economy* 131 (9): 2509–2555. https://doi.org/10.1086/724319. https://doi.org/10.1086/724319.

Foster, Andrew D., and Mark R. Rosenzweig. 2021. "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size." Publisher: The University of Chicago Press, *Journal of Political Economy* 130, nos. 636-680 (October): 000–000. ISSN: 0022-3808, accessed January 18, 2022. https://doi.org/10.1086/717890. http://www.journals.uchicago.edu/doi/full/10.1086/717890.

Fowlie, Meredith, Stephen P. Holland, and Erin T. Mansur. 2012. "What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NOx Trading Program." *American Economic Review* 102, no. 2 (April): 965–93. https://doi.org/10.1257/aer.102.2.965. https://www.aeaweb.org/articles?id=10.1257/aer.102.2.965.

Fowlie, Meredith L. 2009. "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage." *American Economic Journal: Economic Policy* 1, no. 2 (August): 72–112. https://doi.org/10.1257/pol.1.2.72. https://www.aeaweb.org/articles?id=10.1257/pol.1.2.72.

Gillingham, Kenneth, and James H Stock. 2018. "The cost of reducing greenhouse gas emissions." Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2418, *Journal of Economic Perspectives* 32 (4): 53–72.

Gollin, Douglas, and Christopher Udry. 2021. "Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture." _Eprint: https://doi.org/10.1086/711369, *Journal of Political Economy* 129 (1): 1–80. https://doi.org/10.1086/711369. https://doi.org/10.1086/711369.

Harstad, Bård. 2020. *Trade and Trees: How Trade Agreements Can Motivate Conservation Instead of Depletion.* CESifo Working Paper. Publisher: CESifo Working Paper.

———. 2023. "The Conservation Multiplier." Publisher: The University of Chicago Press, *Journal of Political Economy* 131, no. 7 (July): 1731–1771. ISSN: 0022-3808, accessed October 14, 2023. https://doi.org/10.1086/723637. https://doi.org/10.1086/723637.

Harstad, Bård, and Torben K Mideksa. 2017. "Conservation contracts and political regimes." Publisher: Oxford University Press, *The Review of Economic Studies* 84 (4): 1708–1734.

Hausman, Catherine, and Samuel Stolper. 2020. *Inequality, Information Failures, and Air Pollution* [in en]. Technical report w26682. Cambridge, MA: National Bureau of Economic Research, January. Accessed September 18, 2020. https://doi.org/10.3386/w26682. http://www.nber.org/papers/w26682.pdf.

Head, Keith, and Thierry Mayer. 2014. "Gravity equations: Workhorse, toolkit, and cookbook." In *Handbook of International Economics,* 4:131–195. Elsevier.

———. 2023. "Poor Substitutes? Counterfactual Methods in IO and Trade Compared." *The Review of Economics and Statistics* (September): 1–51. ISSN: 0034-6535, accessed November 6, 2023. https://doi.org/10.1162/rest_a_01369. https://doi.org/10.1162/rest_a_01369.

Head, Keith, and John Ries. 2001. "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade." Ratio estimator, *American Economic Review* 91, no. 4 (September): 858–876. https://doi.org/10.1257/aer.91.4.858. https://www.aeaweb.org/articles?id=10.1257/aer.91.4.858.

Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies* 60, no. 3 (July): 497–529. ISSN: 0034-6527, accessed June 25, 2024. https://doi.org/10.2307/2298122. https://doi.org/10.2307/2298122.

Hsiao, Allan. 2022. *Coordination and Commitment in International Climate Action: Evidence from Palm Oil.* Working Paper. February.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." Publisher: MIT Press, *The Quarterly journal of economics* 124 (4): 1403–1448.

IPCC, ed. 2023. "Emissions Trends and Drivers." In *Climate Change 2022 - Mitigation of Climate Change: Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change,* 215–294. Cambridge: Cambridge University Press. ISBN: 978-1-00-915792-6. https://doi.org/10.1017/9781009157926.004. https://www.cambridge.org/core/product/82CFC42906149CC0F91311E9903F9FB8.

Jayachandran, Seema. 2013. "Liquidity Constraints and Deforestation: The Limitations of Payments for Ecosystem Services." *American Economic Review* 103, no. 3 (May): 309–13. https://doi.org/10.1257/aer.103.3.309. https://www.aeaweb.org/articles?id=10.1257/aer.103.3.309.

Jayachandran, Seema, Joost de Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas. 2017. "Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation." _Eprint: https://www.science.org/doi/pdf/10 *Science* 357 (6348): 267–273. https://doi.org/10.1126/science.aan0568. https://www.science.org/doi/abs/10.1126/science.aan0568.

Jenkin, Fleeming. 1872. "On the Principles which regulate the Incidence of Taxes." *Proceedings of the Royal Society of Edinburgh* 7:618–631. https://doi.org/10.1017/S0370164600042760.

Kalouptsidi, Myrto, Paul T Scott, and Eduardo Souza-Rodrigues. 2021. "Identification of Counterfactuals in Dynamic Discrete Choice Models" [in en]. *Quantitative Economics* 12, no. 2 (May): 351–403.

Khandelwal, Amit. 2010. "The Long and Short (of) Quality Ladders." *The Review of Economic Studies* 77, no. 4 (October): 1450–1476. ISSN: 0034-6527. https://doi.org/10.1111/j.1467-937X.2010.00602.x. https://doi.org/10.1111/j.1467-937X.2010.00602.x.

Lagakos, David, and Michael E. Waugh. 2013. "Selection, Agriculture, and Cross-Country Productivity Differences." *American Economic Review* 103, no. 2 (April): 948–980. https://doi.org/10.1257/aer.103.2.948. https://www.aeaweb.org/articles?id=10.1257/aer.103.2.948.

Le Moigne, Mathilde, Simon Lepot, Ralph Ossa, Marcos Ritel, and Dora Simon. 2024. *Greening Ricardo: Environmental Comparative Advantage and the Environmental Gains From Trade.* Working Paper. July.

Nath, Ishan. 2022. *Climate Change, The Food Problem, and the Challenge of Adaptation through Sectoral Reallocation* [in en]. Working Paper.

Perino, Grischa, Robert Ritz, and Arthur van Benthem. 2023. *Overlapping Climate Policies.* Technical report. October.

Phillips, Dom. 2019. "Bolsonaro Declares 'the Amazon Is Ours' and Calls Deforestation Data 'Lies.'" *The Guardian* (July). www.theguardian.com/world/2019/jul/19/jair-bolsonaro-brazil-amazon-rainforest-deforestation.

Rafey, Will. 2023. "Droughts, Deluges, and (River) Diversions: Valuing Market-Based Water Reallocation." *American Economic Review* 113, no. 2 (February): 430–71. https://doi.org/10.1257/aer.20201434. https://www.aeaweb.org/articles?id=10.1257/aer.20201434.

Redding, Stephen, and Anthony J Venables. 2004. "Economic geography and international inequality." Publisher: Elsevier, *Journal of International Economics* 62 (1): 53–82.

Restrepo, Verónica, and Gabriel Leite Mariante. 2024. *Does Conservation Work in General Equilibrium?* Working Paper. January. https://vsalazarr.github.io/jobmarketpaper/jmp_VeronicaSalazarRestrepo_VSR_20240108.pdf.

Sayre, Jay. 2023. *Farm to Firm: Clustering and Returns to Scale in Agricultural Supply Chains.* Working Paper, May. Accessed July 19, 2023.

Scott, Paul T. 2013. *Dynamic Discrete Choice Estimation of Agricultural Land Use* [in en]. Working Paper. Toulouse School of Economics.

Shapiro, Joseph S. 2016. "Trade Costs, CO2, and the Environment." *American Economic Journal: Economic Policy* 8, no. 4 (November): 220–54. https://doi.org/10.1257/pol.20150168. https://www.aeaweb.org/articles?id=10.1257/pol.20150168.

Skiba, Katarzyna. 2023. "Europe's Oldest and Largest Forest Is Now A Major Political Battleground." *WorldCrunch* (August). https://worldcrunch.com/world-affairs/poland-forest.

Song, Xiao-Peng, Matthew C. Hansen, Stephen V. Stehman, Peter V. Potapov, Alexandra Tyukavina, Eric F. Vermote, and John R. Townshend. 2018. "Global land change 1982-2016." *Nature* 560, no. 7720 (August): 639–643. ISSN: 0028-0836, accessed October 5, 2023. https://doi.org/10.1038/s41586-018-0411-9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6366331/.

Sotelo, Sebastian. 2020. "Domestic Trade Frictions and Agriculture." *Journal of Political Economy* 128, no. 7 (June): 2690–2738. Accessed June 16, 2022. https://www.journals.uchicago.edu/doi/abs/10.1086/706859?af=R.

Souza-Rodrigues, Eduardo. 2019. "Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis" [in en]. Publisher: Oxford Academic, *The Review of Economic Studies* 86, no. 6 (November): 2713–2744. ISSN: 0034-6527, accessed September 21, 2020. https://doi.org/10.1093/restud/rdy070. http://academic.oup.com/restud/article/86/6/2713/5232206.

Syverson, Chad. 2011. "What determines productivity?" Publisher: American Economic Association, *Journal of Economic literature* 49 (2): 326–365.

Tombe, Trevor. 2015. "The Missing Food Problem: Trade, Agriculture, and International Productivity Differences." *American Economic Journal: Macroeconomics* 7 (3): 226–58. Accessed May 2, 2023. https://doi.org/10.1257/mac.20130108.

UNFCCC. 2013. *Estimation of carbon stocks and change in carbon stocks of trees and shrubs in A/R CDM project activities.* Technical report v04.1. United Nations Framework Convention on Climate Change, October.

Tables and Figures

Table 1: Summary of key estimated parameters.

Y = Agricultural land use shares: Equation (11)			
Parameter	Description	Value (SE or CI)	Table reference
γ	Deforestation semi-elasticity	$1.43\ (1.39, 1.47)$	2
γ_τ	Travel time semi-elasticity of agriculture share	$-0.02\ (-0.04, -0.005)$	2
ϕ^{FA}	Switching cost semi-elasticity	$-5.35\ (-12.1, -3.26)$	2
Y = Expenditure shares / demand: Equation (15)			
Parameter	Description	Value (SE or CI)	Table reference
σ^A	Trade elasticity within agriculture	9.1 (0.84)	3
σ^M	Trade elasticity within manufacturing	5.5 (1.4)	4

NOTES: Highlights preferred estimates for each key elasticity in the paper. Top panel are land use parameters, $\vec{\gamma}$. Bottom panel are demand-side parameters, $\vec{\sigma}$. Estimating equation references are discussed in Section 4. Table references give details on where the key parameter estimates are located in figures. $Y =$ is used to indicate the left-hand side variable of the key estimating equation.

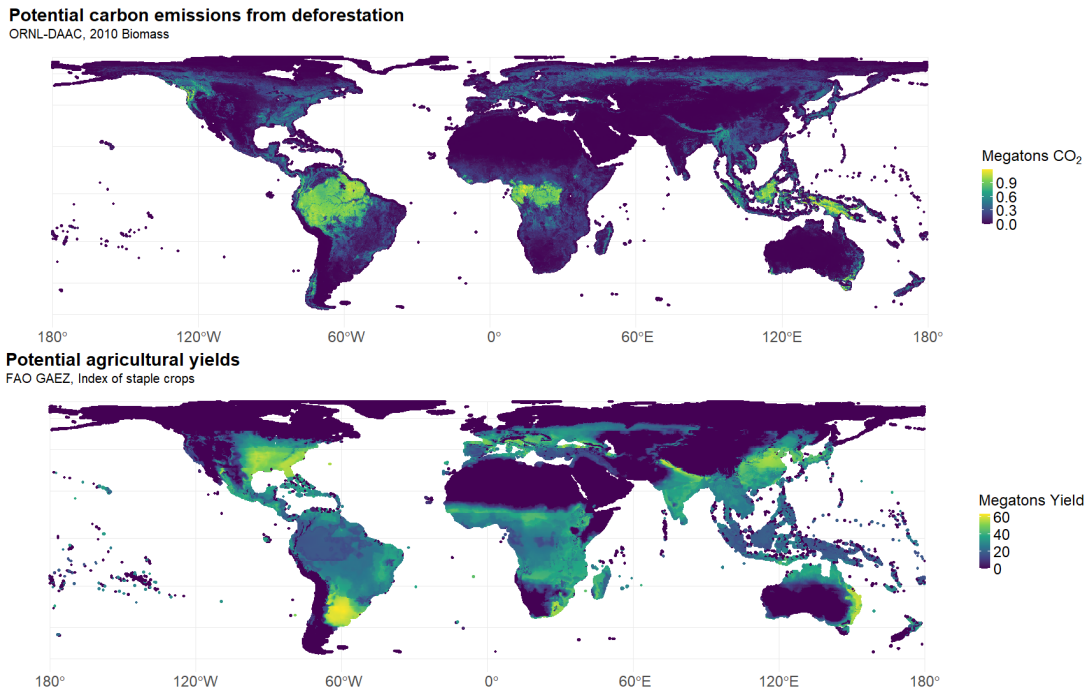


Figure 1: Map of forest biomass (top) and agricultural productivity (bottom), each in megatons.

NOTES: Top figure plots the biomass extracted from the ORNL-DAAC data in 2010. Bottom figure plots the agricultural potential yield index, corresponding to the first principal component of the yield of staple crops discussed in Appendix C. All quantities are rescaled in megatons.

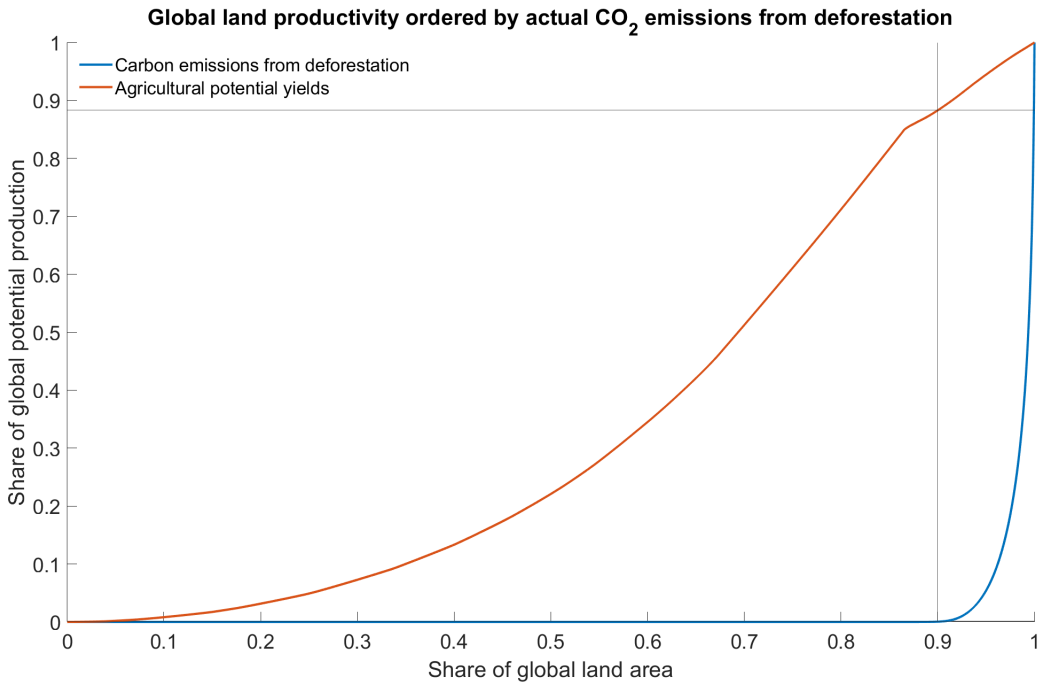


Figure 2: Cumulative share of deforested emissions since 1982 as a share of global land area, compared with the corresponding cumulative share of global yields.

NOTES: Blue line plots the cumulative carbon emissions of all global land area, ordered by realized carbon emissions from deforestation since 1982. Orange line plots the yields of all global land area, ordered by the same emissions. Dashed gray line indicates the 90th percentile.

Table 2: Summarizing estimates of (11).

	(1) OLS	(2) IV	(3) SFD	(4) SFD-IV
γ, revenue	1.119***	1.430***	1.17***	1.05***
	(0.04)	(0.04)	(0.29, 3.99)	(0.53, 2.40)
γ_τ, travel time	-0.02***	-0.02***	0.01	-0.01
	(-0.1 , -0.01)	(-0.04, -0.005)	(0.01,0.58)	(-0.06,0.52)
ϕ^{FA}, lagged forest share	-7.87***	-5.35***	-4.5***	-6.1^*
	(-31.1,-4.21)	(-12.1 , -3.26)	(-15,-1.26)	$(-11.2, 1.79)$
R^2	0.48	0.48		
Observations	4,549,776	4,549,776	4,549,776	4,549,776
country fixed effects	✓	✓	✓	✓
Avg. elasticity equivalent of γ	0.07	0.09	0.07	0.07

NOTES: Regressions implement Equation 11 on the sample of land with an interior choice ($\in (0,1)$) probability of both forest and agriculture, with 687,947 observations per cross-section. Standard errors are computed with 100 iterations of a bootstrap of 5,000 50×50 km blocked clusters sampled with replacement with robust confidence intervals from Efron (1987). Column (1) reflects an OLS estimate. Column (2) includes an instrument for price (WTO accession variety instrument). Columns (3) and (4) are the respective results of a non-linear least squares and GMM spatial differences estimation which implement Equation (12). Reported estimates are both quasi-differenced in the X-direction.

Table 3: Results of second stage of gravity estimation, agricultural industry.

	Levels (1)	First difference (2)	First difference, IV (3)	FD, IV, Controls (4)
Exporter market access	-0.8921***	-0.8979***	-0.8903***	-0.9099***
	(0.0144)	(0.0074)	(0.0107)	(0.0104)
Employment	1.336***			0.8975
	(0.3676)			(0.6692)
Nitrogen fertilizer	0.1718**			-0.0296
	(0.0835)			(0.0318)
Labor share	-2.380***			-0.0752
	(0.3026)			(0.2168)
Pesticide	0.3229***			0.0025
	(0.1029)			(0.0546)
R^2	0.87450	0.86549	0.86543	0.88296
Observations	3,935	4,300	4,300	3,691
F-test (1st stage), Exporter market access			1,709.7	1,465.8
$\hat{\sigma}$	9.266	9.794	9.117	11.10
Std. Err., $\hat{\sigma}$	1.170	0.6702	0.8407	1.227
Exporter fixed effects	✓	✓	✓	✓

NOTES: Implements the exporter second-stage regression (16) for the agricultural industry, $h = A$, using the ITPD-E trade flow database, macroeconomic data from the Penn World Table, FAOSTAT data, and IFASTAT proprietary data. All controls are expressed in logs in column (1) and in first differences of logs in (4). Standard errors are clustered at the exporter level to allow for serial correlation. Reported standard errors for $\hat{\sigma}$ are delta method standard errors. Instruments in (3) and (4) are lags of leave-one-out market access.

Table 4: Results of second stage of gravity estimation, manufacturing industry.

	Levels (1)	First difference (2)	First difference, IV (3)	FD, IV, Controls (4)
Exporter market access	-0.6986***	-0.7906***	-0.8183***	-0.8074***
	(0.0178)	(0.0098)	(0.0526)	(0.0543)
Employment	1.265**			0.9365***
	(0.5907)			(0.2922)
R^2	0.97529	0.70960	0.70879	0.71204
Observations	4,409	4,254	4,254	4,114
F-test (1st stage), Exporter market access			31.631	30.499
$\hat{\sigma}$	3.318	4.777	5.504	5.192
Std. Err., $\hat{\sigma}$	0.1634	0.1980	1.442	1.316
Exporter fixed effects	✓	✓	✓	✓
year × Exporter	✓			

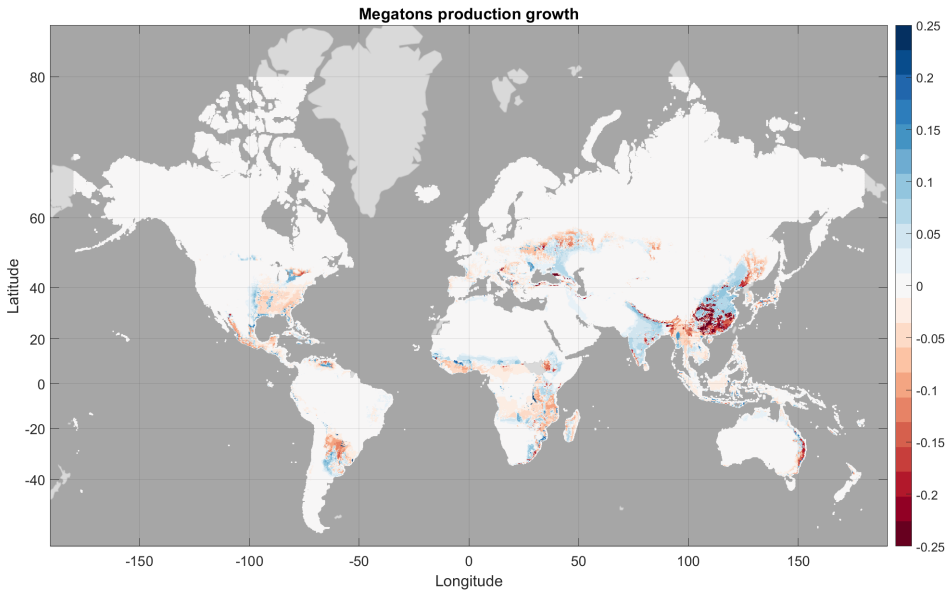
NOTES: Implements the exporter second-stage regression (16) for the manufacturing industry, $h = M$, using the ITPD-E trade trade flow database, macroeconomic data from the Penn World Table, FAOSTAT data, and IFASTAT proprietary data. All controls are expressed in logs in column (1) and in first differences of logs in (4). Standard errors are clustered at the exporter level to allow for serial correlation. Reported standard errors for $\hat{\sigma}$ are delta method standard errors. Instruments in (3) and (4) are lags of leave-one-out market access.

Table 5: Results for global Pigouvian tax of $190 under various demand-side assumptions.

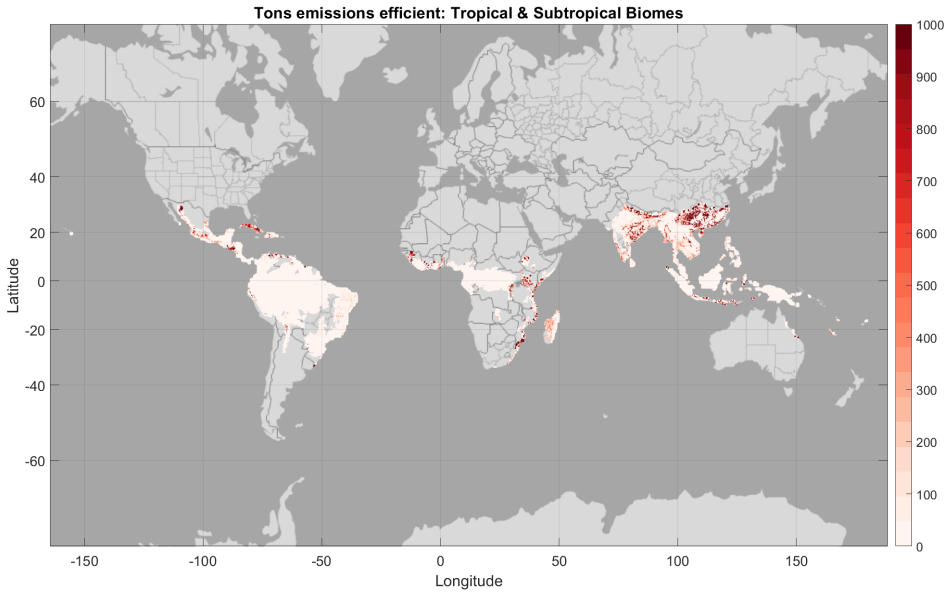
Outer nest σ	$\sigma = 1$		$\sigma = 0.5$	
Inner nest	$\sigma^A = 4$	$\sigma^A = 9.1$	$\sigma^A = 4$	$\sigma^A = 9.1$
Emissions	0.03	0.03	0.04	0.03
(prop. of BAU)	(0.03, 0.03)	(0.03, 0.03)	(0.03, 0.04)	(0.03, 0.04)
Agriculture	0.88	0.87	0.94	0.93
(prop. of BAU)	(0.86, 0.89)	(0.85, 0.87)	(0.92, 0.94)	(0.92, 0.94)
Manufacturing	1.02	1.01	1.01	1.01
(prop. of BAU)	(1.01, 1.02)	(1.01, 1.02)	(1.01, 1.02)	(1.01, 1.02)
Cropland area	0.84	0.84	0.86	0.87
(prop. of BAU)	(0.84, 0.84)	(0.84, 0.84)	(0.86, 0.86)	(0.87, 0.87)
Welfare	0.95	0.97	0.94	0.96
(prop. of BAU)	(0.95, 0.96)	(0.97, 0.97)	(0.94, 0.95)	(0.95, 0.96)
Prices	1.77	1.64	2.25	2.17
(PTR, %)	(1.66, 1.90)	(1.56, 1.73)	(2.10, 2.42)	(2.06, 2.30)
Wages	0.19	0.13	0.26	0.20
(PTR, %)	(0.16, 0.23)	(0.12, 0.16)	(0.23, 0.32)	(0.18, 0.23)

NOTES: Prop. of BAU indicates the change in quantity, in proportional terms, to the business-as-usual equilibrium ($\hat{Q} = Q^{TAX}/Q^{BAU}$). PTR % indicates pass-through as a percent of the $190 tax rate. Standard errors indicate bias-corrected bootstrap draws (Efron 1987) with 100 draws. Each draw pulls from the sampling distribution of $(\vec{\gamma}, \sigma^M)$ reported in Table 1. Columns (2) and (4) use the estimated value and confidence interval of σ^A from Table 1, while Columns (1) and (3) use $\sigma^A = 4$, a calibrated value . The outer nest σ has no confidence interval attached, as it as a calibrated value. Estimates are produced for the values of θ_i^A calibrated off of World Bank International Comparison Program data. I report a welfare-weighted average across countries $PTR = \sum_{i=1}^{J} W_i^{BAU}(p_i^{TAX} - p_i^{BAU})/(\sum_{i=1}^{J} W_i^{BAU})$. Cropland area refers to the share of global land area in agriculture, $\frac{1}{N}\sum_{\Omega} \mu^A(\omega; s_0(\omega))$.

Figure 3: Geography of agricultural production and deforestation-induced carbon dioxide emissions after a $190/ton Pigouvian tax.



(a) Changes in agricultural production (megatons yield)



(b) Remaining emissions after the Pigouvian tax

NOTES: In the top panel, I plot differences of agricultural production under a Pigouvian tax of $190 on deforestation-induced carbon emissions relative to the business-as-usual scenario, $\eta^A(\omega)[\mu^{A,TAX} - \mu^{A,BAU}]$. Agricultural yields are measured in megatons of an aggregate crop. In the bottom panel, I plot the remaining emissions in the taxed equilibrium, focusing on the tropical biomes of the world, $d(\omega)\mu^{A,BAU}$. For visual clarity, I rescale these emissions into tons. I focus on tropical biomes as classified by maps from the World Wide Fund for Nature: these are Tropical and Subtropical Coniferous Forests, Dry Broadleaf Forests, Grasslands, Savannas and Shrublands, and Moist Broadleaf Forests.

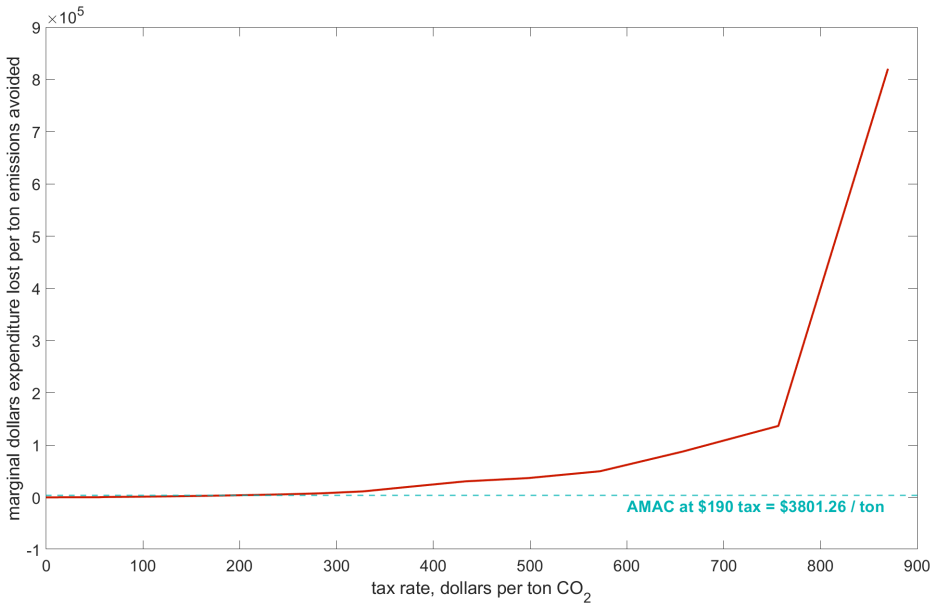


Figure 4: Global average marginal abatement cost curve, $AMAC$.

NOTES: Simulates the Pigouvian tax counterfactual for a logarithmically spaced grid of tax rates (social costs of carbon) from $[10^{-3}, 10^{3}]$. Simulations assume parameters (γ, σ) are at the point estimates from Table 1. Plots the change in total global equivalent variation per additional ton abated, $\Delta EV/\Delta D$, discarding the first grid point.

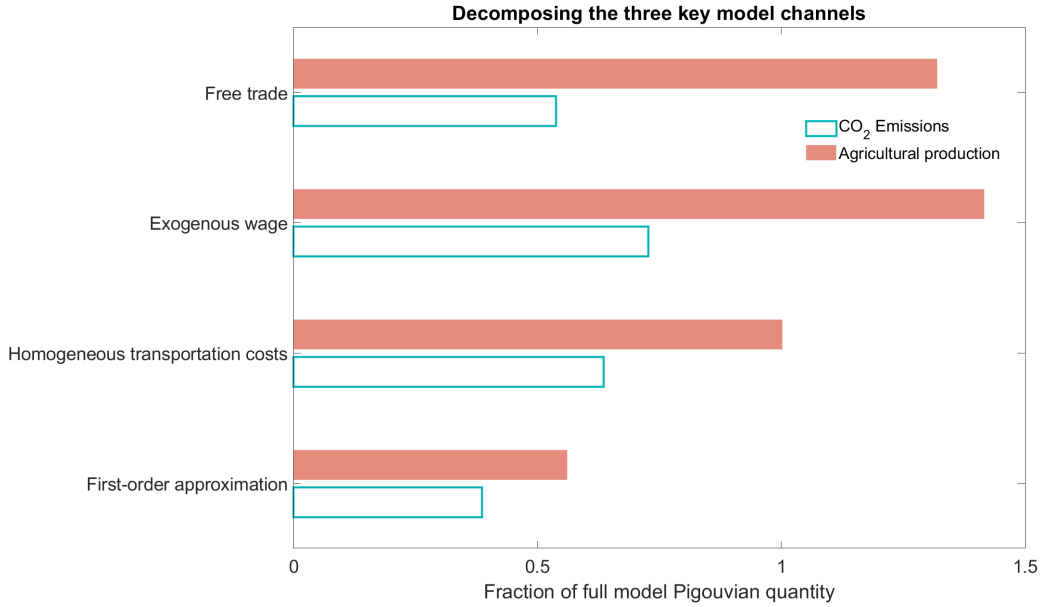


Figure 5: Decomposition: CO_2 emissions and agricultural production under alternate model assumptions which shut down key mechanisms.

NOTES: Resimulates model at a \$190 Pigouvian tax under various assumptions. Bottom bars indicate a fully first-order approximation where prices and wages are both fixed at values from the baseline equilibrium. Top three labels respectively simulate: all trade barriers are set to 0, wages fixed at baseline equilibrium values, and transportation costs are set to the within-country 10th percentile. Blue un-filled bars indicate carbon dioxide emissions under each scenario as a fraction of the value in the full model Pigouvian tax. Red filled bars indicate agricultural production.

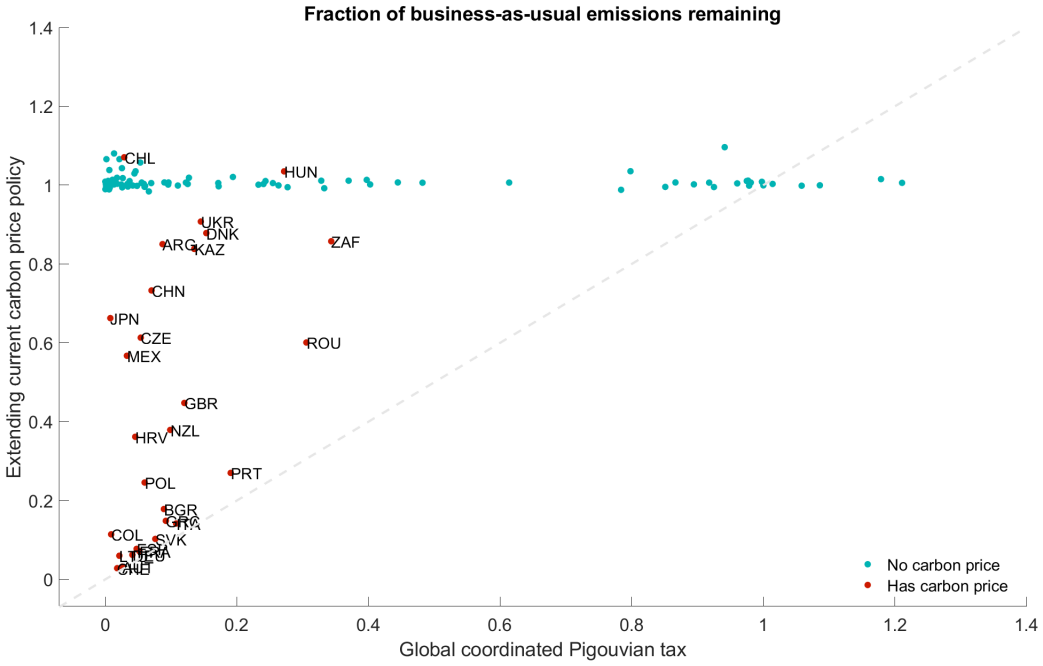


Figure 6: Share of BAU deforestation emissions abated under maximum existing ETS and carbon tax rates on deforestation-related carbon emissions.

NOTES: Indicates the average abatement cost per country based on Equation (9). Tax rates and ETS prices are from the World Bank's Carbon Pricing Dashboard. Current carbon prices, on the y-axis, indicate the share of BAU emissions removed under these prices, with a negative share corresponding to an increase in deforestation-related emissions. The x-axis plots shares of BAU abated under the global tax rate. The gray line indicates the 45 degree line, so that countries along this line abate similar amounts under either pricing regime

Appendix: The Global Allocative Efficiency of Deforestation

A Theoretical appendix

A.1 Aggregate income

This proof follows directly from Section 3.5 of Train (2009). Total surplus accrued to the landowner is given by the maximum rent they can charge. Leveraging the properties of the extreme value distribution, and under the assumption that the error $\epsilon^h(\omega)$ is i.i.d. with scale parameter $\frac{1}{\gamma}$, one can write the money-metric surplus as follows.

$$\begin{aligned}\Pi &= \sum_{\omega=1}^{N} \mathbb{E} \max_{h \in \{F,A\}} \pi_\epsilon^h(\omega) \\ &= \frac{1}{\gamma}\left(\aleph + \sum_{\Omega} \log \sum_{h \in \{F,A\}} \exp[\gamma \pi^h(\omega)]\right)\end{aligned}$$

A.2 Existence of an equilibrium

To derive the existence of an equilibrium for my market equilibrium construct, I first re-write the market equilibrium system from equations (6) and (7) in terms of excess demand equations. I then use an application of Brouwer's theorem to demonstrate existence.

Market clearing gives rise to a candidate excess demand function, scaled by total landowner production Q^h. In agriculture, there is an excess demand function for each country.

$$Z_j^A(\vec{p}) = \frac{1}{p_j Q_j^A}[E_j^A(p) - p_j Q_j^A(p)] = 0$$

In manufacturing, the analogous excess demand function is defined in terms of labor supply. This labor supply term $M_i = \sum_{\omega=1}^{N_i} \mu_\epsilon^F(\omega; s_0)$ comes out of the discrete choice problem of the landowner. The greater the volume of labor which is free from agriculture, the greater the share of manufacturing labor:

$$Z^M(\vec{p}) = \frac{1}{\sum_{j=1}^{J} w_j M_j}[E_j^M(p) - w_j M_j(p)] = 0$$

The excess demand system constitutes $2 \times N - 1$ nonlinear equations in $2 \times N$ unknown prices. It admits a fixed point under a normalization, given continuity of the supply and demand functions.

Define a transition function to update prices as the product of prices and a term which scales with excess demand according to a computational parameter $\nu \in (0, 1)$.

$$\mathbf{T}(\vec{p}) = \vec{p}[1 + \nu Z(\vec{p})]$$

I show that this procedure has a fixed point $T(\vec{p}) = \vec{p}$ by way of Brouwer's fixed point theorem.

The first condition for existence is continuity of Z. This can be verified through continuity of demand and supply, each. Demand (X) is continuous in prices. (the Armington CES admits continuity as a fractional polynomial of price). Supply (Q^A, M) is also continuous under the assumption that $\epsilon^h(\omega)$ is distributed Type I extreme value (or any choice probability which yields interior solutoins for ω, Kalouptsidi, Scott, and Souza-Rodrigues 2021). Under this assumption, the choice probabilities μ are given by the inverse logistic function in profits, which is quasilinear in price.

It remains to show the excess demand function Z is bounded below by -1 so an iteration results in nonnegative prices. At any price, expenditure $E(p)$ is bounded below by 0. Then, at $E(p) = 0$, note that $Z(p)$ is by construction exactly -1. Then, the transformation $\mathbf{T}$ along with a normalization to appease Walras's Law maps a convex set onto itself, $[-1, \infty)^{2\times(N-1)} \to [-1, \infty)^{2\times(N-1)}$. Brouwer's theorem ensures the existence of a fixed point.

A.3 Planner's problem

The social planner maximizes a weighted sum of country-level representative utility functions. Each utility function U_i is weighted by Pareto weights λ_i. For notational simplicity, I rewrite the utility function representative consumer in i from Equation (3) as $V_i(X_i^A, X_i^M) - \mu D$, where $V_i(X_i^A, X_i^M)$ corresponds to the nested CES utility over manufacturing X_i^M and agricultural goods X_i^M.[1] Then, the social welfare function, with Pareto weights λ_i (these sum to 1 without loss), is,

1. In this section, the representative consumer in each country has the same absolute disutility of emissions μ, but in general they need not: see Kotchen and Powers (2006).

$$\max_{X,D,\mu^A} -\mu D + \sum_{i=1}^{J} \lambda_i V_i(X_i^A, X_i^M)$$

$$\text{s.t.} \quad \sum_{j=1}^{J} T_{ij} X_{ij}^A = Q_i^A = \sum_{\omega=1}^{N_i} \eta^A(\omega)\mu^A(\omega; s_0(\omega)) \quad \forall i \in \{1, 2, \ldots, J\}$$

$$\sum_{j=1}^{J} T_{ij} X_{ij}^M = Q_i^M = \bar{\eta}_i^M M_i \quad \forall i \in \{1, 2, \ldots, J\}$$

$$D = \sum_{\omega=1}^{N_i} d(\omega)\mu^A(\omega; s_0(\omega))$$

$$\mu^A(\omega; s_0(\omega)) + \mu^F(\omega; s_0(\omega)) \leq 1 \quad \forall \omega \in \Omega \quad \cdots \text{Feasibility constraint}$$

$$\sum_{i=1}^{J} X_{ij}^M = X_i^M$$

$$\sum_{i=1}^{J} X_{ij}^A = X_i^A$$

Here, the first two constraints determine constrain total consumption of each country i-sector h-level good to equal production of that good. The production of the externality D is tied directly to the production function for agriculture in the next constraint. Define the multiplier on each of these three constraints as $\kappa_i^A, \kappa_i^M, \kappa^D$, respectively.

The next set of constraints are feasibility constraint on land, where land use shares in each plot ω must sum to 1. The constraint on plot ω is assigned a multiplier $\kappa(\omega)$, representing the shadow value (rental rate) of land of type ω. Importantly, the planner's solution is conditionally efficient: it takes as given the starting allocation of land uses $s_0(\omega)$. Finally, the last two constraints connect country-level demand to aggregate sectoral demand.

Thus, the planner's solution differs from market clearing in (6) and (7) because the planner solves a public goods coordination problem. In the market, no one consumer has control over D (they contribute on average $\frac{1}{N}D$ to total emissions, for N large). Then, they do not individually rationally abate (Bergstrom, Blume, and Varian 1986). Note that unlike in some traditional public goods problems, where the public good is produced by a centralized entity, my setting mirrors work by Chichilnisky and Heal (1994) where emissions are produced in a decentralized fashion. In contrast, the planner sets land use shares directly and overcomes this coordination problem.

Motivating the Pigouvian tax. I next consider the marginal cost of abatement, κ^D, which is the cost of change emissions by a marginal amount (the multiplier on the emissions

constraint). To derive this, I first take the system's first-order conditions with respect to X and D. These are:

$$\begin{aligned}\lambda_j \frac{\partial V_j}{\partial X_{ij}^A} &= T_{ij}\kappa_i^A \\ \lambda_j \frac{\partial V_j}{\partial X_{ij}^M} &= T_{ij}\kappa_i^M \\ -\mu &= \kappa^D\end{aligned}$$

The third first-order condition is crucial for the Pigouvian tax. The social planner sets marginal abatement costs κ^D to exactly offset the disutility from global emissions μ. Because these first-order conditions hold for every X_{ij}, they hold for X_{ii}. Without loss trade barriers with oneself are normalized to 1 (non-unit trade barriers should show up in within-country transportation costs, so this is without loss), $T_{ii} = 1$, so this recasts the first-order condition such that trade barriers do not directly enter:

$$\begin{aligned}\lambda_i \frac{\partial V_i}{\partial X_{ii}^A} &= \kappa_i^A \\ \lambda_i \frac{\partial V_i}{\partial X_{ii}^M} &= \kappa_i^M \\ -\mu &= \kappa^D\end{aligned}$$

With these identities in hand, I can now derive an explicit analytic marginal abatement cost. For any plot ω, the marginal abatement cost can be derived by taking the first order condition with respect to agricultural land shares $\mu^A(\omega; s_0(\omega))$:

$$\kappa^D = \frac{1}{d(\omega)}\left(\eta^A(\omega)\kappa_i^A - \bar{\eta}_i^M a_i \kappa_i^M\right) \quad \{\forall \omega : d(\omega) \neq 0\}$$

That is, plot-level marginal abatement costs are equalized across land which has forest ($d(\omega) \neq 0$). Abatement costs on ω are determined by three empirical objects: emissions intensity $d(\omega)$, agricultural productivity, and the size of the labor externality in i coming from land reallocation.

Using the first order conditions on X_i^A, X_i^M derived in the first step, I obtain an important result:

$$\kappa^D = \frac{\lambda_i}{d(\omega)}\left(\eta^A(\omega)\frac{\partial V_i}{\partial X_{ii}^A} - \bar{\eta}_i^M a_i \frac{\partial V_i}{\partial X_{ii}^M}\right) \quad \{\forall \omega : d(\omega) \neq 0\}$$

The marginal abatement cost chosen depends on the marginal utility of consumption in each country (Chichilnisky and Heal 1994). This result highlights the importance of calculating differences in real welfare across countries as in Equation (9).

Finally, combining this last marginal abatement cost equation with the first-order-condition on D, the social planner sets marginal abatement costs κ^D to be the disutility from emissions in this quasilinear setting $\mu = \kappa^D$. Thus, a Pigouvian tax on emissions $t^{SCC}d(\omega) = \kappa^D d(\omega) = \mu d(\omega)$ implements the planner's solution in the full market equilibrium.

More on damages. This utility cost μ (equivalently, damages) of a unit of emissions today corresponds to the foregone future consumption of agriculture and manufacturing. Future consumption can be lost through several mechanisms: a common modeling choice is to reflect climate change as a loss in productivity in levels. Importantly, the true Pigouvian tax $t = \kappa^D$ is clearly function of current GDP and emissions. As a result, the tax itself is potentially endogenous variable and should respond in equilibrium. Previous work abstracts from this concern by looking at truly marginal changes in emissions. In my setting, given that deforestation is 11% of global carbon emissions, an effective carbon tax can have an order of magnitude effect on atmospheric CO_2. In sensitivity checks, I allow for the social cost of carbon (the optimal tax rate) to endogenously respond to remaining atmospheric emissions.

Sensitivity: manufacturing emissions. In a further sensitivity check, I consider the case where the social planner also attends to emissions in the non-agricultural sector. I consider a simple linear emissions technology in manufacturing. Given the average CO_2 emissions per unit $\bar{d}_i^M$, total emissions in country i are given by:

$$D_i = \bar{d}_i^M \bar{\eta}_i^M M_i + \sum_{\omega=1}^{N_i} d(\omega)\mu^A(\omega; s_0(\omega))$$

Then, the social planner must also contend with a new spillover channel; increasing manufacturing labor supply M_i will drive up emissions in the manufacturing sector. Accounting for this channel changes the first order condition with respect to $\mu^A(\omega; s_0(\omega))$ to be:

$$\kappa^D = \frac{\lambda_i}{d(\omega) - \bar{d}_i^M a_i \bar{\eta}_i^M} \left(\eta^A(\omega) \frac{\partial V_i}{\partial X_{ii}^A} - \bar{\eta}_i^M a_i \frac{\partial V_i}{\partial X_{ii}^M} \right) \quad \{\forall \omega : d(\omega) \neq 0\}$$

The social planner will discourage deforestation only when the spillover effects onto manufacturing in country i are small enough: if a plot emits less than its spillover onto manufacturing $d(\omega) < \bar{d}_i^M a_i \bar{\eta}_i^M$, the social planner prefers to deforest it.

A.4 Comparative statics and identification

This section derives comparative statics for an equilibrium with a Pigouvian tax of t on agricultural deforestation. This comparative statics exercise clarifies how the model is identified. The key challenge in this model relative to conventional trade models, which are identified in relative changes (Dekle, Eaton, and Kortum 2007), is that the supply-side is not identified in multiplicative changes (Kalouptsidi, Scott, and Souza-Rodrigues 2021). I discuss the key assumptions required to maintain identification.

Market clearing in the counterfactual equilibrium. The initial market clearing conditions, copying equations (7) and (6) are:

$$p_i Q_i^A = \sum_{j=1}^{J} E_{ij}^A$$
$$w_i M_i = \sum_{j=1}^{J} E_{ij}^M$$

These same conditions must hold in the new equilibrium. Denote the new equilibrium outcomes by a prime.

$$p_i' Q_i^{A'} = \sum_{j=1}^{J} E_{ij}^{A'}$$
$$w_i' M_i' = \sum_{j=1}^{J} E_{ij}^{M'}$$

Denote by the hat $\hat{x} = \frac{x'}{x}$ the ratio of the value of x in the new ($'$) and old equilibria. Beginning with the first equilibrium condition, I divide both sides by the starting value of production in agriculture to obtain hat changes:

$$\begin{aligned}\hat{p}_i\hat{Q}_i^A = \frac{1}{p_i Q_i^A} p_i' Q_i^{A'} &= \sum_{j=1}^{J} \frac{1}{p_i Q_i^A} E_{ij}^{A'} \\ &= \sum_{j=1}^{J} \frac{X_{ij}^A}{p_i Q_i^A} \hat{E}_{ij}^A \\ &= \sum_{j=1}^{J} \gamma_{ij}^A \hat{E}_{ij}^A\end{aligned}$$

In the last line, γ_{ij}^A denotes the sales share of j in i's agricultural production in the *baseline equilibrium*. I can repeat this same exercise for the manufacturing clearing condition.

$$\hat{w}_i \hat{M}_i = \sum_{j=1}^{J} \gamma_{ij}^M \hat{E}_{ij}^M$$

The hat change in the expenditure $\hat{E}_{ij}^h$ is composed of three terms:

$$\hat{E}_{ij}^h = \hat{\lambda}_j^h \hat{\lambda}_j^{i|h} \hat{Y}_j$$

Hat-changes in outer nest shares are:

$$\hat{\lambda}_j^h = \left(\frac{P_j^{h'}}{P_j^{C'}}\right)^{1-\sigma} \left(\frac{P_j^h}{P_j^C}\right)^{\sigma-1} = \left(\frac{\hat{P}_j^h}{\sum_{h\in\{A,M\}} \lambda_j^h (\hat{P}_j^h)^{1-\sigma}}\right)^{1-\sigma}$$

By a similar token, hat-changes in inner nest shares are:

$$\hat{\lambda}_j^{i|h} = \left(\frac{p_i^{h'}}{P_j^{h'}}\right)^{1-\sigma^h} \left(\frac{p_i^h}{P_j^h}\right)^{\sigma^h-1} = \left(\frac{\hat{p}_i^h}{\sum_{k=1}^{J} \lambda_j^{k|h} (\hat{p}_k^h)^{1-^h}}\right)^{1-\sigma^h}$$

Finally, the change in income can be written as a weighted sum of changes in landowner income and wage worker income, where I define α_j as the share of income in country j from landowners in business-as-usual $\alpha_j = \frac{\Pi_j}{Y_j}$. Without explicit cross-country labor mobility or population growth, there is no change in the total potential labor supply L_j:

$$\hat{Y}_j = \alpha_j \hat{\Pi}_j + (1 - \alpha_j)\hat{w}_j$$

Supply-side identification. Then, a final ingredient in identification focuses on changes in land use shares (an input into $\hat{M}_i$ and $\hat{Q}_i^A$) and land rents, $\hat{\Pi}_j$. To demonstrate identification, I will assume that the error term ϵ^h in the landowners problem, Equation (1), is distributed as a Type-I Extreme value: several other functional form assumptions will yield similar identification results (Chiong, Galichon, and Shum 2016). The variance of the shock is represented by an additional parameter γ. Under this assumption, land use shares on plot ω aggregate over parcel-level shocks $\epsilon^h(\omega)$ as:

$$\mu^A(\omega; s_0(\omega)) = \frac{\exp[\gamma \pi^A(\omega; s_0(\omega))]}{\exp[\gamma \pi^A(\omega; s_0(\omega))] + \exp[\gamma \pi^F(\omega; s_0(\omega))]}$$

Then, profits can be linked directly to (interior) land use shares as:

$$\log \frac{\mu^A(\omega; s_0(\omega), 0)}{\mu^F(\omega; s_0(\omega), 0)} = \gamma[\pi^A(\omega; s_0(\omega), 0) - \pi^F(\omega; s_0(\omega), 0)]$$

This same expression will hold for counterfactual land use shares and profits:

$$\log \frac{\mu^A(\omega; s_0(\omega), t^{SCC})}{\mu^F(\omega; s_0(\omega), t^{SCC})} = \gamma[\pi^A(\omega; s_0(\omega), t^{SCC}) - \pi^F(\omega; s_0(\omega), t^{SCC})]$$

Then, plugging in profit functions, some knowledge of the level of initial prices and wages is required to solve for a counterfactual allocation:

$$\log \frac{\hat{\mu}^A}{\hat{\mu}^F} = \gamma[(p_i' - p_i)\eta^A(\omega) - a_i(w_i' - w_i) - t^{SCC} d(\omega)]$$

Conditional on identifying $\hat{\mu}$, agricultural production changes $\hat{Q}_i^A$, labor supply changes $\hat{M}_i$, and profit changes $\hat{Pi}_i$ all follow. Changes in agricultural quantities are:

$$\hat{Q}_i = \frac{\sum_{\omega=1}^{N_i} \eta^A(\omega)\mu^A(\omega; s_0(\omega), t^{SCC})}{\sum_{\omega=1}^{N_i} \eta^A(\omega)\mu^A(\omega; s_0(\omega), 0)}$$

The hat-change in manufacturing labor supply is given by:

$$\hat{M}_i \frac{\sum_{\omega=1}^{N_i}[1-\mu^A(\omega; s_0(\omega), t^{SCC})}{\sum_{\omega=1}^{N_i}[1-\mu^A(\omega; s_0(\omega), 0)]}$$

and finally, changes in aggregate profits are given by the level change in the inclusive value (where $\aleph$ is the Euler-Mascheroni constant, see Appendix A.1):

$$\hat{\Pi}_i = \frac{\aleph + \sum_{\omega=1}^{N_i} \log[\exp(\gamma\pi^A(\omega; s_0(\omega), t^{SCC})) + \exp(\gamma\pi^F(\omega; s_0(\omega), t^{SCC}))]}{\aleph + \sum_{\omega=1}^{N_i} \log[\exp(\gamma\pi^A(\omega; s_0(\omega), 0)) + \exp(\gamma\pi^F(\omega; s_0(\omega), 0))]}$$

A key identification assumption. I highlight a key assumption required to maintain identification. Changes in wages must be identified in levels for landowners, and in relative changes for workers. Due to the normalization of one wage from Walras's Law), the *level* of wages for landowners is pinned down by the shifter a_i, which I later calibrate to match the labor share in agriculture. Thus, as the above equation elucidates, if a_i is fixed, changes in wage levels are also identified (in the language of Kalouptsidi, Scott, and Souza-Rodrigues (2021), a_i converts wage ratios $\hat{w}_i$ into pre-specified additive changes). This shifter represents labor-specific productivity in agricultural production. If labor-specific productivity does not change, wages identified in changes can be converted to level changes. Otherwise, I must specify an endogenous a_i to maintain identification.

Identification also maintains a more classical assumption from the discrete choice literature: the outside option does not systematically shift across land. Relative to discrete choice models, my use of the aggregate price index to calculate welfare partially endogenizes the value of the outside option. Where this assumption would fail then comes from cross-plot, within-country shifts in land rents, which might for example come from unmodeled rent dynamics from urban land.

Summary. The counterfactual system can be solved without a normalization as there are $2 \times N$ conditions in the $2 \times N$ unknowns $(\hat{p}, \hat{w})$. Changes are purely identified off of the parameters $(\gamma, \sigma, \sigma^A, \sigma^M)$, so that other land use parameters in Equation (1) do not enter counterfactuals directly (save through predicting baseline prices p_i and wages w_i).

Appendix Table A.1 provides an accounting of the key moments which provide identification of the carbon tax counterfactual and their sources. I note the absence of two moments which one could match. First, I could match the exact levels of land use $\mu^A(\omega; s_0(\omega))$ by calibrating a residual profit shifter. I do not do this in my preferred approach because these shifters have ambiguous interpretation for final misallocation: they can represent measurement error, or other market and policy distortions. Without further detail, such distortions

unclear implications for the normative allocative efficiency of deforestation. Second, I could set residual bilateral preferences θ_{ij}^h to exactly match cross-country flows. However, I lack data on flows for several key countries and thus prefer to use an imputed estimate discussed in the main text to avoid overfitting.

Table A.1: List of parameters and the moments which they match.

Moment	Description	Parameter	Source
	Land use parameters		
$\mathbb{E}\left[\frac{\mu^A(\omega;s_0(\omega))}{\mu^F(\omega;s_0(\omega))}\vert\eta^A(\omega)\right]$	Land use response to yields	γ	Estimate
$\mathbb{E}\left[\frac{\mu^A(\omega;s_0(\omega))}{\mu^F(\omega;s_0(\omega))}\vert s_0(\omega)\right]$	Inertia in land use change	ϕ^{FA}	Estimate
$\frac{w_i(L_i-M_i)}{\Pi_i}$	Labor share of ag. value add	a_i	Calibrate
	Demand-side parameters		
$\mathbb{E}\left[\frac{d\log\lambda_{ij}^h}{d\log p_i^h}\right]$	Cross-country share elasticity	$1-\sigma^h$	Estimate
$\mathbb{E}\left[\frac{d\log\lambda_j^h}{d\log P_i^h}\right]$	Cross-sector share elasticity	$1-\sigma$	Calibrate, 0.5
λ_j^i	Sectoral expenditure shares	θ_j^h	Calibrate
λ_{ij}^h	Cross-country expenditure shares	T_{ij}	Estimate
	Welfare calculations		
$\frac{\Pi_i}{Y_i}$	Agriculture share of value-add	α_i^{LAND}	Calibrate
$\frac{\Pi_J}{P_J^C}$	Level of agricultural value add	w_J	Calibrate

A.5 Nested CES expenditure shares

This derivation proceeds in three steps. I first derive demand in the inner nest, holding fixed total sectoral expenditure in the outer nest. In order to get outer nest expenditure shares, I derive an intermediate step which redefines the outer nest problem as a function of the solution to the inner nest problem. Finally, I calculate outer nest demand.

Step 1: obtaining inner nest demand. Beginning with inner nest demand, define the total expenditure on goods in sector h, the solution to the outer nest, as $E_i^h = \sum_{i=1}^J E_{ij}^h$. For readability, I introduce a piece of notation: define the sector- and country pair-specific price p_{ij}^h which equals $p_{ij}^A = p_i T_{ij}$ in agriculture and $p_{ij}^M = p_i^M T_{ij}$ in manufacturing. Holding fixed outer nest expenditures, the consumer problem is as follows.

$$\max_{X_{ij}^h} \left(\sum_{i=1}^{J} \left(\theta_{ij}^h X_{ij}^h \right)^{\frac{\sigma^h - 1}{\sigma^h}} \right)^{\frac{\sigma^h}{\sigma^h - 1}}$$
$$\text{s.t.} \sum_{i=1}^{J} p_{ij}^h X_{ij}^h = E_j^h$$

To get to a demand function, I start with the first-order condition with respect to X_{ij}^h. Define a multiplier on the budget constraint ν:

$$\left(\sum_{i=1}^{J} \left(\theta_{ij}^h X_{ij}^h \right)^{\frac{\sigma^h - 1}{\sigma^h}} \right)^{\frac{1}{1-\sigma^h}} \left((\theta_{ij}^h)^{\sigma^h - 1} X_{ij}^h \right)^{-\frac{1}{\sigma^h}} = \nu p_{ij}^h$$

Taking a ratio of these first-order conditions for two country's goods, X_{ij}^h and X_{nj}^h, gives the classic Armington result that relative demand is pinned down by relative prices up to a preference parameter (Anderson 1979).

$$\frac{X_{ij}^h}{X_{nj}^h} = \frac{(\theta_{nj}^h)^{\sigma^h - 1} (p_{ij}^h)^{-\sigma^h}}{(\theta_{ij}^h)^{\sigma^h - 1} (p_{nj}^h)^{-\sigma^h}}$$

Rearranging this ratio and aggregating across the J country-specific first-order conditions pins down the total sector-level expenditure E_j^h in terms of the expenditure on any one origin country's good, E_{nj}^h:

$$X_{ij}^h = \frac{(\theta_{nj}^h)^{\sigma^h - 1} (p_{ij}^h)^{-\sigma^h}}{(\theta_{ij}^h)^{\sigma^h - 1} (p_{nj}^h)^{-\sigma^h}} X_{nj}^h$$
$$\therefore E_j^h \coloneqq \sum_{i=1}^{J} p_{ij}^h X_{ij}^h = \sum_{i=1}^{J} \frac{(\theta_{nj}^h)^{\sigma^h - 1} (p_{ij}^h)^{1-\sigma^h}}{(\theta_{ij}^h)^{\sigma^h - 1} (p_{nj}^h)^{1-\sigma^h}} E_{nj}^h$$

At last, I can now use this result to derive cross-country expenditure shares, which can be used to pin down demand. For the manufacturing industry, I substitute the endogenous wage w_i into p_{ij}^h using Equation (2). The representative consumer spends share $\lambda_j^{i|M} = \frac{E_{ij}^M}{E_j^M}$ of total manufacturing expenditures E_j^M on goods from i:

$$\lambda_j^{i|M} = \frac{(\theta_{ij}^M \bar{\eta}_i^M)^{\sigma^M - 1}(T_{ij} w_i)^{1-\sigma^M}}{\sum_{n=1}^J (\theta_{nj}^M \bar{\eta}_n^M)^{\sigma^M - 1}(T_{nj} w_n)^{1-\sigma^M}} \tag{A.1}$$

Agricultural prices are not analytically linked to wages, so the final expression depends directly on prices. The representative consumer spends share $\lambda_j^{i|A} = \frac{E_{ij}^A}{E_j^A}$ of total agricultural expenditures E_j^A on goods from i:

$$\lambda_j^{i|A} = \frac{(\theta_{ij}^A)^{\sigma^A - 1}(T_{ij} p_i)^{1-\sigma^A}}{\sum_{n=1}^J (\theta_{nj}^A)^{\sigma^A - 1}(T_{nj} p_n)^{1-\sigma^A}}$$

Given cross-country expenditure shares, demand is recovered using prices $X_{ij}^h = \frac{E_{ij}^h}{p_{ij}^h} = \frac{\lambda_j^{i|h}}{p_{ij}^h} E_j^h$. Then, the inner nest problem is solved as a function of outer nest expenditures E_j^h.

Step 2: reformulating outer nest problem. To get demand from the outer nest, I now use the inner nest demand function to re-write the outer nest consumer choice problem. To do this, I derive a within-sector price index across countries, which is the cost of purchasing an additional utile of h in j. Substituting expenditure shares into the inner nest utility function,

$$U_j^h = \left(\sum_{i=1}^N (\theta_{ij}^h)^{\frac{\sigma^h - 1}{\sigma^h}} \left(\overbrace{\frac{\lambda_j^{i|h}}{p_{ij}^h} E_j^h}^{X_{ij}^h} \right)^{\frac{\sigma^h - 1}{\sigma^h}} \right)^{\frac{\sigma^h}{\sigma^h - 1}} = E_j^h \left(\sum_{n=1}^J (\theta_{nj}^h)^{\sigma^h - 1}(p_{ij}^h)^{1-\sigma^h} \right)^{\frac{1}{1-\sigma^h}}$$

The second equality is purely algebraic. Inverting this gives the unit cost of E_j^h, which is exactly the Dixit-Stiglitz price index.

$$P_j^h = \left(\sum_{n=1}^J (\theta_{nj}^h)^{\sigma^h - 1}(p_{nj}^h)^{1-\sigma^h} \right)^{\frac{1}{1-\sigma^h}}$$

Thus, the solution to the outer nest trades off these sector-level price indices. To see this, recast the aggregate utility function as an implicit utility function with a single nest:

$$U_j(x) = \left(\sum_{h \in \{A,M\}} (U_j^h)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} - \mu D$$
$$\text{s.t. } P_j^M U_j^M + P_j^A U_j^A = Y_j$$

Step 3: obtaining outer nest demand. The problem is now reduced to an equivalent problem to the inner nest. Derivation of the final step, the outer nest expenditure shares, will thus follow the same algebra used in the first-order conditions for the previous inner nest demand shares. Given that total expenditure across goods will exactly equal total income in i, we can use the above preference relation to derive expenditure shares on each sector:

$$\lambda_i^h = \begin{cases} \frac{(\theta_i^h)^{\sigma-1}(P_i^h)^{1-\sigma}}{\sum_{h \in \{A,M\}} (\theta_i^h)^{\sigma-1}(P_i^h)^{1-\sigma}} & \text{if } \sigma \in (0,1) \cap (1,\infty) \\ \theta_i^h & \text{if } \sigma = 1 \end{cases} \tag{A.2}$$

where I single out the Cobb-Douglas case.

A.6 Comparative statics for nested CES

In this section I present the elasticity of the share of expenditure shares for the representative consumer in country j spent on agriculture from i, λ_{ij}^A, to agricultural prices in i, p_i. Expenditure shares λ_{ij}^A can be written in the following differential form:

$$\frac{d \log \lambda_{ij}^A}{d \log p_i} = (1 - \sigma^A) + (\sigma^A - \sigma) \frac{d \log P_j^A}{d \log p_i} + (1 - \sigma) \frac{d \log P_j^C}{d \log p_i}$$

Two primitives are required to clean up the above expression: derivatives of the price index of agriculture P_j^A and the aggregate consumption price index P_j^C with respect to a change in price in i. These are:

$$\frac{d \log P_j^A}{d \log p_i} = \left(\frac{P_j^A}{T_{ij} p_i} \right)^{\sigma^A - 1} = \lambda_j^{i|A}$$
$$\frac{d \log P_j^C}{d \log p_i} = \frac{d \log P_j^C}{d \log P_j^A} \cdot \frac{d \log P_j^A}{d \log p_i} = \left[\left(\frac{P_j^C \theta_j^A}{P_j^A} \right)^{\sigma - 1} \cdot \frac{d \log P_j^A}{d \log p_i} \right] = \lambda_j^A \cdot \lambda_j^{i|A}$$

Then, substituting these back into the original differential yields the price elasticity of the

expenditure share.

$$\begin{aligned}\frac{d\log\lambda_{ij}^{A}}{d\log p_i} &= (1-\sigma^{A}) + (\sigma^{A}-\sigma)\frac{d\log P_j^{A}}{d\log p_i} + (1-\sigma)\frac{d\log P_j^{C}}{d\log p_i} \\ &= (1-\sigma^{A}) + (\sigma^{A}-\sigma)\lambda_j^{i|A} + (1-\sigma)[\lambda_j^{A}\lambda_j^{i|A}] \\ &= \underbrace{(1-\sigma^{A})}_{\substack{\text{single-level CES} \\ \text{(classic Armington)}}} + \underbrace{\lambda_j^{i|A}[\sigma^{A}-\sigma+(1-\sigma)\lambda_j^{A}]}_{\text{propagation to outer nest}}\end{aligned}$$

I note two edge cases of the aggregate demand share elasticity:

- when $\lambda_j^{i|A} \to 0$, so i is not a key agricultural supplier for j, consumers in j will not respond to shocks in i.
- when $\lambda_j^{i|A} \to 1$, the elasticity reduces to $(1-\sigma)(1+\lambda_j^{A})$, which is a positive number for $\sigma \in (0,1)$. Put differently, as a country i becomes a near-monopolist over agriculture in j, small changes in i lead to a *positive* change in the expenditure share.
- In the limiting case of the Cobb-Douglas preference between manufacturing and agriculture ($\sigma \to 1$), expenditure share elasticities approach 0.

The aggregate expenditure elasticity follows from an application of the product rule. It includes an income elasticity:

$$\frac{d\log E_{ij}}{d\log p_i} = \left(\frac{d\log\lambda_{ij}^{A}}{d\log p_i} - 1\right) + \frac{d\log Y_j}{d\log p_i}$$

This income effect term is small for small open economies i when $j \neq i$. When $j = i$, income effects depend on inframarginal firm profits. See:

$$\begin{aligned}\frac{d\log Y_i}{d\log p_i} &= \frac{1}{Y_i}\left(\frac{d\Pi_i}{dp_i} + w_i\frac{dM_i}{dp_i}\right) \\ &= \frac{1}{Y_i}(Q_i - \gamma a_i w_i N_i \mathbb{E}_i[\eta^{A}(\omega)\tilde{s}(\omega)]) \\ &= \frac{1}{Y_i}\sum_{i=1}^{\omega} q^{A}(\omega)[1-\gamma a_i w_i \mu^{F}(\omega; s_0(\omega))]\end{aligned}$$

where $\tilde{s}(\omega)$ is the agricultural price-agricultural land use share elasticity as described in the main text Subsection 3.5. Without a manufacturing industry (that is, $p_i Q_i = Y_i$ agriculture describes the entire economy and wages are thus 0 without loss, $w_i = 0$), the income effect is

entirely pinned down by $Q_i/Y_i = p_i$. Relaxing this restriction, the size of the income effect decreases as manufacturing shares in income rise.

A.7 Pass-through in general equilibrium

Recall from the main text that the general equilibrium goods market clearing conditions imply the following identities for a change in the Pigouvian tax rate t.

$$\frac{dQ_i^A}{dp_i}\frac{dp_i}{dt} + \frac{dQ_i^A}{dw_i}\frac{dw_i}{dt} + \frac{dQ_i^A}{dt} = \sum_{j=1}^{J}(1+T_{ij})\left[\frac{dX_{ij}^A}{dp_i}\frac{dp_i}{dt} + \frac{dX_{ij}^A}{dw_i}\frac{dw_i}{dt}\right]$$
$$\frac{dM_i}{dp_i}\frac{dp_i}{dt} + \frac{dM_i}{dw_i}\frac{dw_i}{dt} + \frac{dM_i}{dt} = \sum_{j=1}^{J}(1+T_{ij})\left[\frac{dX_{ij}^M}{dp_i}\frac{dp_i}{dt} + \frac{dX_{ij}^M}{dw_i}\frac{dw_i}{dt}\right]$$

Algebraically manipulating the first equation gives the following pass-through formula for prices:

$$\frac{dp_i}{dt} = \frac{\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dw_i}\frac{dw_i}{dt} - \frac{dQ_i^A}{dw_i}\frac{dw_i}{dt} - \frac{dQ_i^A}{dt}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dp_i} + \frac{dQ_i^A}{dp_i}}$$

Immediately, this is exactly the partial equilibrium pass-through rate *less* a (potentially negative) adjustment factor. The adjustment reflects the fact that market wages will absorb some of the tax incidence:

$$\frac{dp_i}{dt} = \overbrace{\frac{-\frac{dQ_i^A}{dt}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dp_i} + \frac{dQ_i^A}{dp_i}}}^{\text{Partial equilibrium pass-through } \text{PTR}_p^{\text{PE}}} - \underbrace{\frac{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dw_i} + \frac{dQ_i^A}{dw_i}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dp_i} + \frac{dQ_i^A}{dp_i}}\frac{dw_i}{dt}}_{\text{General equilibrium adjustment}}$$

The wage pass-through equation implies a similar relationship:

$$\frac{dw_i}{dt} = \overbrace{\frac{-\frac{dM_i}{dt}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^M}{dw_i} + \frac{dM_i}{dw_i}}}^{\text{Partial equilibrium pass-through } \text{PTR}_w^{\text{PE}}} - \frac{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^M}{dp_i} + \frac{dM_i}{dp_i}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^M}{dw_i} + \frac{dM_i}{dw_i}}\frac{dp_i}{dt}$$

Solving for the pass-through of the tax onto prices gives that the general equilibrium pass-through is a linear transformation of partial-equilibrium pass-through rates for prices PTR_p^{PE} in terms of the price and wage elasticities:

$$\frac{dp_i}{dt} = \frac{1}{1-a_1}(\text{PTR}_p^{\text{PE}} - a_2)$$

where the coefficient a_1 depends on the ratio of elasticities in each industry. For example, if the agriculture sector as a whole (meaning in terms of the sum of demand and supply elasticities) is relatively more elastic than manufacturing to *wages*, then a_1 is larger and the overall pass-through onto prices is lower:

$$a_1 = \frac{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dw_i} + \frac{dQ_i^A}{dw_i}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dp_i} + \frac{dQ_i^A}{dp_i}} \times \frac{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^M}{dp_i} + \frac{dM_i}{dp_i}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^M}{dw_i} + \frac{dM_i}{dw_i}}$$

Then, the affine adjustment a_2 represents a relative downward shift in price-specific pass-through due to the partial-equilibrium pass-through for wages, PTR_w^{PE}:

$$a_2 = \frac{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dw_i} + \frac{dQ_i^A}{dw_i}}{-\sum_{j=1}^{J}(1+T_{ij})\frac{dX_{ij}^A}{dp_i} + \frac{dQ_i^A}{dp_i}}\text{PTR}_w^{\text{PE}}$$

I conclude this derivation with two observations. First, general equilibrium pass-through as implied by the model is a linear combination of pass-through rates when only wages or prices respond. Second, the general equilibrium price pass-through rate can either be larger or smaller than the partial equilibrium pass-through rate. Without a price channel, wages are falling (because taxes raised manufacturing labor supply $\frac{dM_i}{dt} > 0$), putting downward pressure on pass-through onto prices (because a_2 will in turn be a positive number). However, if one believes in the "regular" case that cross-price elasticities are smaller than own-price elasticities, then $a_1 < 1$, creating a second "multiplier" effect on the partial-equilibrium pass-through. The combination of these two adjustments is ambiguous.

B Model estimation appendix

B.1 Calibration appendix

Calibration involves setting labor shifters a_i such that the prevailing wage w_i in country i matches the labor share of agricultural value-added in FAO data. For $\sigma = 0.5$, I also calibrate θ_i^A, holding $\theta_i^M = 1$, to exactly match expenditure shares in the World Bank International Comparison Program (for the Cobb-Douglas case, $\sigma = 1$, θ_i^h is set to exactly the shares in the data). I match both datasets in 2000.

On both dimensions, several countries lack data: the Democratic Republic of the Congo importantly has limited macroeconomic data despite having a large tropical forest. In these cases, I impute missing consumption and value-added shares. If the GDP of the country is known in the Penn World Tables, I target the average value of the moment in the same quartile of GDP as the country with missing data. Otherwise, I use the average value in the bottom 25% of GDP. In future work, I plan to use a cross validation-based method to fill in these data gaps (Carleton, Crews, and Nath 2023).

In a robustness check regarding distortions from agricultural policy, I also set values of country-level, additive profit shifters ξ_i, to match land use shares in the data.

I calibrate as follows:

1. Set $\vec{\gamma}$, $\vec{\sigma}$ to their estimated values in Table 1.

2. Guess a vector of prices and wages, $\vec{p}$

3. If setting ξ_i, solve the following nonlinear equation based on (1) to set the average share of agriculture in i to its value in the data. Else, move to next step.

$$\frac{1}{N_i}\sum_{\Omega_i}\frac{\exp(\pi^A(\omega)+\xi_i)}{\exp(\pi^A(\omega)+\xi_i)+1} = \text{Share agriculture in } i \text{ in 2000}$$

4. Iterate over a_i to match value-added shares, based on Equation (1):

$$a_i = \frac{\text{Value Added Data}_i \times \Pi_i}{w_i \times (L_i - a_i M_i)}$$

5. Set θ_i^h so that, at $\vec{p}$, the predicted share of expenditure on agriculture λ_i^A in Equation (A.2) matches the World Bank International Comparison data, 2001 vintage

$$\theta_i : \lambda_i^A = \text{ICP Share spent on agriculture}_i$$

6. Check whether wages and prices satisfy goods market clearing conditions, Equations (6) and (7). If so, the procedure ends; if not, update $\vec{p}$.

Given the values of $\vec{\xi}_i$, $\vec{a}_i$ (and, for $\sigma = 0.5$, $\vec{\theta}_i$) which are calibrated from this procedure, I then run the model with the new starting land use from the 2000 data (which was perfectly matched by the model). I seek a vector of prices and wages which would comprise an equilibrium from this 2000 land use and compare it with the data in 2016. Thus, the calibration procedure targets moments in the year 2000, but leaves moments in the year 2016 un-targeted. Prices and wages are identified in the calibration procedure by finding the transfer price which would set Q_i^h, which is implied directly by land use data in my framework in Section 3.3, to aggregate demand, which is again coming from the data.

Other values in the model are calibrated directly using the data. I set average manufacturing productivities $\bar{\eta}_i^M$ to the value-added per worker in manufacturing, given by the total GDP of the country, less agricultural value added, over manufacturing employment. This method is unable to provide a TFP measure for some countries: I impute these countries with the average non-missing TFP.

I adopt a simplified labor supply: total labor supply is endowed one-to-one with land $L_i = N_i$. As a_i accounts for technological differences across countries, this assumption matters most within-country. It assumes the only driver of variance in the wage bill across farms is output productivity, not labor-specific productivity. I thus understate differences in the wage bill across, for example, distant and near-to-market land.

B.2 Calculating the carbon tax equivalent of an arbitrary distortion

For an arbitrary wedge – trade costs, spatially heterogeneous production costs, or cross-country differences in non-agricultural productivity – I consider the effective change in the cost of abating the marginal unit of emissions induced by that wedge. Call the wedge of interest w. I index the baseline model with a superscript BAU, and the equilibrium without a given wedge with superscript w. For the three wedges of interest, I specifically solve for the following counterfactual equilibria.

1. Trade costs: set all trade costs T_{ij} to theoretical free trade, $T_{ij} = 1$
2. Heterogeneous production costs: set all transportation costs within a given country to the 10th percentile of costs within that country
3. Differences in non-agricultural productivity: set TFP in every country to the TFP of the US and hold wages fixed in the Counterfactual equilibrium

The equilibrium emissions from deforestation in country i under these counterfactuals is D_i^w. To calculate the contribution of each mechanism to deforestation, I solve for the market equilibrium in each scenario with the added constraint that emissions in the counterfactual equal emissions in business-as-usual.

$$D_i^w = D_i^{BAU}$$

Call the multiplier on this constraint p_i^w, the shadow price of carbon in i induced by the wedge w. Larger values of p_i^w indicate that the marginal ton of carbon emissions is more valuable to social welfare. To solve for this shadow price, I solve for a market equilibrium vector of prices p_i^w on the landowners' problem,

$$\pi(\omega; s_0(\omega)) = \mathbb{E}_\epsilon \max_{h \in \{F,A\}} \pi_\epsilon^h(\omega; s_0(\omega), p_i^w)$$

I can do this because the social planner (Appendix Section A.3) will always set the marginal abatement cost across all plots of land exactly equal. In this case, the planner is constrained by the emissions level D_i^{BAU}, which pins a unique marginal abatement cost (exactly as in a cap-and-trade system). I solve for p_i^w as a nested contraction mapping. In the outer loop I solve for the market equilibrium through the tatonnement procedure described in Appendix Section A.2, and in the inner loop I solve for p_i^w using the contraction mapping $\mathbf{T}$:

$$\mathbf{T}p_i^w = p_i^w - \log D_i^w(p_i^w) + \log D_i^{BAU}(p_i^w)$$

B.3 Pass-through decomposition

Appendix Figure A.1 presents the bivariate relationship between the spatial correction from Equation (10) and the model-implied pass-through rate at a tax rate of \$190. The two are significantly positively correlated. A 1% increase in the spatial correction term indicates a 0.06% increase in the pass-through rate onto agricultural prices, or around 0.8 basis points of additional pass through. A linear relationship between the two explains 71% of the cross-country variation in pass-through rates (the R^2 of the line of best fit for Appendix Figure A.1).

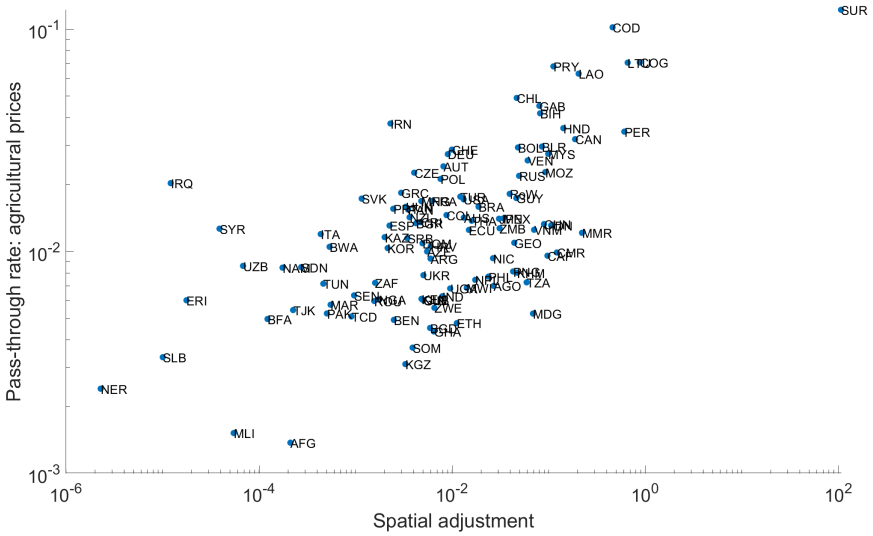


Figure A.1: Plotting the spatial adjustment in Equation (10) against equilibrium pass-through at a $190 tax

NOTES: Plots the spatial adjustment term in Equation (10) and the general equilibrium pass-through rate $(p_i^{TAX} - p_i^{BAU})/t$. Each point is a country. Excludes Egypt, Denmark, Great Britain, and Libya: see Data Appendix C. Axes are in log scales: maximum of y-axis is a pass-through rate of 0.1.

B.4 Comparison of model-implied acreage price elasticities and the literature

Comparing the change in agricultural land area with the change in prices in Table 5, the global area-weighted average acreage-price elasticity of agricultural goods is 0.4. This exercise uses the relationship $\frac{d\log Q(p)}{d\log t} = \frac{d\log Q(p)}{d\log p}\frac{d\log p}{d\log t}$ and inverts it, calculating

$$\frac{d\log Q(p)}{d\log p} = \frac{d\log Q(p)}{d\log t}\left(\frac{d\log p}{d\log t}\right)^{-1}$$

Intuitively, it partials out the change in quantities arising purely from tax-induced price changes relative to the tax instrument. It thus isolates the last link in the causal chain $dt \to dQ \to dp \to dQ$. One can interpret this as a long-run elasticity of deforestation, as it calculates equilibrium responses of otherwise myopic landowners.

My implied acreage-price elasticities aligns with prior work in country-specific contexts. Scott (2013) identifies a long-run acreage-price elasticity in the US of 0.25-0.65: my average estimated US-specific elasticity is 0.5 $(0.22, 0.7)$. Brazil has a higher acreage elasticity of 0.82 $(0.43, 0.97)$, which mirrors prior discussion of an elasticity of 0.64 (Araujo, Costa, and Sant'Anna 2020). Acreage-price elasticities from the USDA for wheat tend to also fall in the

range of 0.33 – 0.4 (Source). Note that these numbers are not strictly comparable, as my figures are generated by a tax instrument, while these papers are identifying off of a price shock: but, their comparable magnitudes should be seen as evidence that the model's final output figures in levels are comparable to peers' policy experiments.

B.5 Agricultural distortions

In my main results, the model calibration does not directly match country-level land use shares (see Appendix Table A.23 for details on targeted moments). To test sensitivity to this moment, I calibrate country-level land quality shifters ξ_i (see Appendix B.1 for details). Calibrating these wedges in agricultural profits, I find that all are positive: agricultural land is underpredicted in business-as-usual if I do not directly match this moment. Thus, I interpret these calibrated values as country-level policies or conditions which improve agricultural land returns. Consistent with prior work, my statistical measure of land market distortions is higher in higher income countries – particularly in the US, EU, and China, which have agricultural subsidy programs larger than the GDP of most high-forest tropical countries (Adamopoulos and Restuccia 2022).[2]

Ex ante, the effect of agricultural subsidies on avoidable emissions is ambiguous. Subsidizing agricultural land would weakly expand agricultural land use: thus, in levels, there should be less forest area overall. However, for emissions reductions, the spatial distribution of subsidies also matters. If subsidies favor low emissions markets more than high emissions markets, the net effect could lower deforestation.

To separate the level effect from this allocative effect, I re-run the Pigouvian tax counterfactual under a model which explicitly accounts for these calibrated country-level shifters. With land market distortions, the Pigouvian tax prevents fewer emissions: 92% (89, 94) of business-as-usual emissions are inefficient. The land area impacted is also larger: 30% of cropland area is reallocated out of agriculture. However, the absolute *level* of remaining agricultural land after the Pigouvian tax exceeds business-as-usual cropland area.

Thus, agricultural subsidy policy expands agricultural production on land which is high-emissions and low-yield. Though tropical regions subsidize agriculture less than non-tropical regions, the overall level effect of subsidies dominates any reallocation of agriculture towards high-subsidy countries. On net, agriculture in the tropics expands.

2. In principle, if these distortions represent subsidies, I could directly endogenize them as a redistributive subsidy from workers to landowners. Instead, I take them as exogenous features of land.

B.6 Further welfare discussions

The discussion in section 6.3 refers to Appendix Table A.2. For each decomposition, I assume that the calibrated values of θ_i and the wage normalization w_J are equal to their status quo values (cf., re-calibrating both to match their respective moments under each set of alternate assumptions).

Table A.2: Equivalent variation (EV) decomposition across key model mechanisms.

Scenario name	Scenario assumption	EV per ton, \$/ton
Baseline	$t = \$190$	\$106 (\$100, \$123)
	Decomposition, $t = \$190$	
Free trade	$T_{ij} = 1$	\$65 (\$55, \$92)
Autarky	$T_{ij} = 999$	\$19 (\$18, \$33)
Homogeneous costs	$c(\omega) = \text{Quantile}_i^{10}(c(\omega))$	\$96 (\$87, \$113)
Wages held fixed	$w_i' = w_i$	\$19 (\$12, \$57)*

NOTES: Each computes equivalent variation per ton of emissions reductions from Equation (9). Equivalent variation excludes countries with no GDP in the Penn World Table, importantly including the Democratic Republic of the Congo. $*$ indicates that the wage-fixed model only uses equivalent variation among landowners, for which the relevant baseline comparison is \$89.

I also provide an additional decomposition in this section: I set, in every country i, productivity in terms of both yields $\eta^A(\omega)$ and emissions $d(\omega)$ to the national average everywhere. This decomposition isolates the component of the deadweight loss of the Pigouvian tax which is driven by heterogeneity across land. I plot the results at the country level in Appendix Figure A.2. The color of each bubble corresponds to the correlation between yields and emissions in the true data. Clearly, there is visual separation. Shutting down within-country correlation leads to higher estimated welfare costs in countries with positive correlation (high yields coincide with high emissions) and lower estimated welfare costs elsewhere. Several large forest basins experience an extremely large miscalculation from the country-level productivity approach: Bolivia, the Congo, Mexico, and Peru stand out as having understated costs absent within-country data. Meanwhile, ignoring heterogeneity dramatically overstates costs in South and Central American countries including Suriname, Gabon, and Nicaragua.

Next, I test the sensitivity of these welfare figures to a variety of more adversarial assumptions.

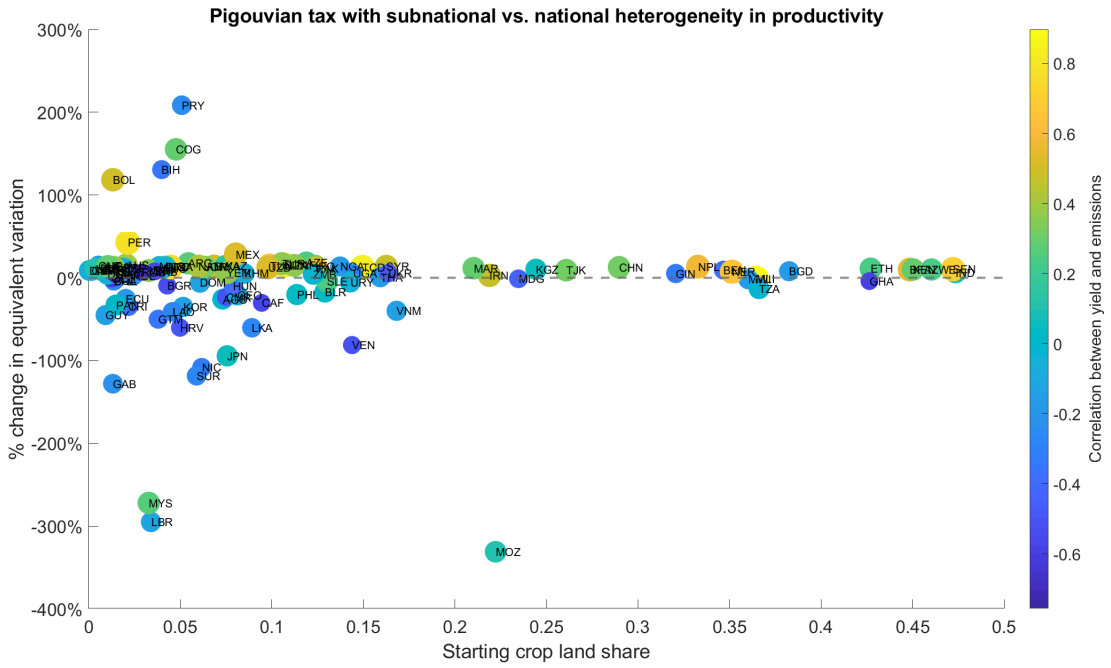


Figure A.2: Change in equivalent variation per ton comparing a Pigouvian tax in a model with only national average yields and emissions with status quo.

NOTES: Positive values indicate larger equivalent variation in no subnational variation scenario. Excludes 13 countries with no GDP in the Penn World Table, importantly including the Democratic Republic of the Congo, as well as outliers Myanmar, Honduras, and Cote d'Ivoire. Correlations indicate a Pearson correlation coefficient.

C Data appendix

In this section, I describe the process of collating a mix of remotely sensing and economic data into a format conducive to training a global high-resolution model. The goal of the data collection process is to obtain a panel $t \in \{1995, \ldots, 2020\}$ of observations at the level of an individual plot, ω.

C.1 Sketch of data construction process

There are multiple types of data which must be synthesized to complete this project. This includes remote-sensing data, the bulk of which is presented in raster (point-cloud) format, administrative data on boundaries, roadways, and waterways which is presented in a vector (geometric) format, and economic data which is often tied to an administrative unit.

To harmonize all of these data sources, I define plots using a rectangular grid overlaid on the world map. Importantly, the raster data comes at varying resolutions, which means that my final dataset is subject to aggregation bias if my resolution is too low, and can create an uninformative block-diagonal data structure if the resolution is too high. Table A.21 lists key raster data sources and describes their resolution.

Grid cells belong to the geographic unit which captures most of their area. The data is then extracted in stages. Each stage is one of 504 tiles which are 10 degrees-by-10 degrees. Doing this ensures that the entire world is extracted in a timely manner.[3]

My final dataset creates grid cells at the scale of 11.1 square kilometers near the equator (0.1 arc-degrees). I define neighborhoods of cells for later clustering and analysis of spillovers. Neighborhoods are 121 square kilometers in area (1 degree), so 100 grid cells form a neighborhood. As is displayed in A.21, most raster datasets have resolution of at least 0.1 degrees. Every plot is a square grid cell, a varying fraction of which is actual usable land.

Importantly, this approach does drop some area from the global dataset. Rather than storing multiple countries' worth of data for every grid cell, I only extract data for the largest country in that grid cell. Generally, this under-represents small countries or countries with particularly irregular boundaries.

C.2 Land use data.

Several satellite-based land use products targeted at measuring deforestation have emerged in recent years. I will rely on two such sources described below: the first, the European Space

3. An alternative is extracting data by country. This avoids having to assign individual plots to a specific country, for example. However, it can lead to multiple observations for a fixed plot of land, which complicates the analysis from the model. I opt for a consistent resolution.

Agency ESACCI data, is used to estimate the model, and the second, the VCF project, is used to validate the model on long-run deforestation data. Additionally, for benchmarking my data, I compare their predictions to long-run country-level data from the Food and Agriculture Organization's 5-yearly Forest Resources Assessment in Appendix Section C.2.1.

First, the European Space Agency Climate Change Initiative, or ESACCI data, which covers a 1 kilometer $\times$ 1 kilometer grid annually from 1995-2019. The ESACCI data is valuable for its high temporal frequency and detailed landcover classification (see A.20 for details on the classification scheme), which I can aggregate into five land uses: forest, settlement, cropland, vegetation, and bare soil/other. However, ESACCI changes measurement regimes in 2016, making the last 5 years of the data spuriously different from the first 20 years. Other work has found that these last 5 years of data dramatically over-predict tropical deforestation. I demonstrate that the first 20 years of the data 1995-2015 align with aggregate FAO statistics in Appendix C.2.1. I primarily use the ESACCI data as a left-hand side variable for my empirical strategy in identifying the landowners' problem in Section 4.

Second, I use the vegetation continuous fields product. This "VCF" measure is taken at a 250 meter resolution derived from the MODIS satellite (Song et al. 2018). Prior work has used this data in the context of Indonesian palm oil (Hsiao 2022). It reports land shares of tree canopy, short vegetation, and bare areas. Unfortunately, short vegetation is not mapped one-to-one with cropland; it will also include grasslands and meadows. I have two cross-sections of this data, one indicating the level of tree cover in 2016, and one indicating the change in tree cover since 1982. I use VCF for the construction of my model, validating results from ESACCI.

Given both data are classification-based machine learning products (assigning continuous satellite imagery a discrete land use label), they necessarily induce non-classical measurement error (Alix-Garcia and Millimet 2022; Torchiana et al. 2022). This measurement error correction is thus econometrically important, and I intend to explore corrections in future work. Pilot tests of the correction suggested by Torchiana et al. (2022) on ESACCI data indicate that central land use elasticity estimates are robust in South America.

C.2.1 Comparing remote sensing and national accounts data

I compare my land use data with aggregate statistics, such as those published by the FAO, and attempt to detail exactly where the two diverge. Deviations tend to occur in countries with very little land in forest or agriculture, which motivates removing such countries from the sample. Most of these countries are smaller island nations.

Appendix Figure A.3 shows the difference between the total, country-level (log) land area in hectares devoted to agriculture. Points above the 45-degree line in this figure generally

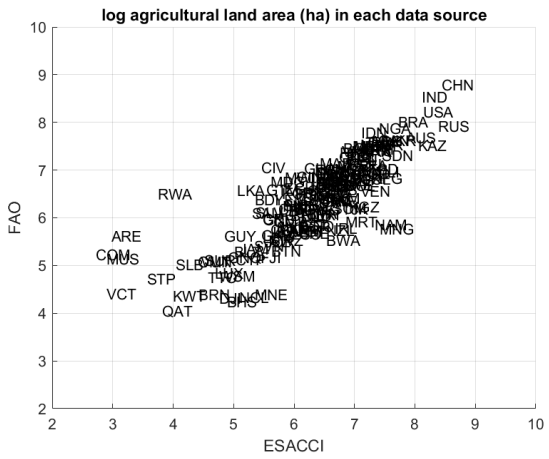


Figure A.3: Compares volume agricultural land in national accounts data (FAO) and remote sensing data (ESACCI)

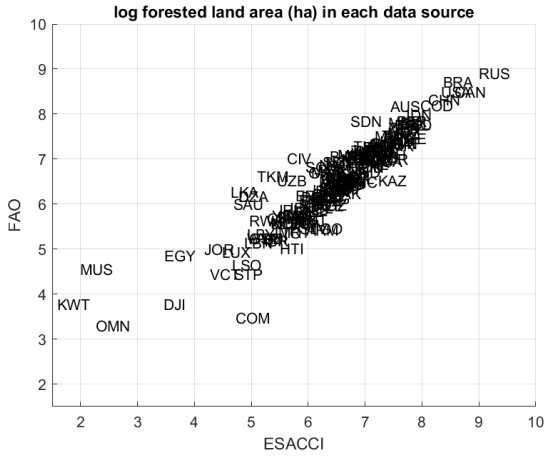


Figure A.4: Compares volume forest land in national accounts data (FAO) and remote sensing data (ESACCI)

indicate the satellite data underestimates the total agricultural area relative to the FAO data: underestimation is an issue in the lower range of the data. As stated before, aggregation is a significant potential reason for under-estimation in this lower range, relative to ground-based counts. The ESACCI algorithm sensing will pick up on agriculture only when it dominates a 9 square kilometer grid cell. This aggregation story is most clear in the severe underprediction of agriculture in irregularly shaped (often island) nations such as Mauritius or Comoros on the left-hand extreme of the plot.

The FAO Forest Resource Assessments allow me to compare remotely sensed with national accounts of forested land use. I repeat the exercise of validation with the landcover data against the FRA data in Appendix Figure A.4. Once again, I find that nations with very

low forest cover tend to have unreliable measurement in the remote sensing data. However, on average, in the cross section, I recover a consistent measure of forest cover with these national accounts data.

C.3 Productivity data.

C.3.1 Emissions and biomass.

In the main text, I describe how I convert biomass (a measure of how much carbon is stored in trees) into emissions from deforestation. I use two sources of data on biomass. First, I use a single global cross-sectional map from NASA ORNL-DAAC which reports aboveground and belowground biomass in 2010. Data is provided at a 30 meter-by-30 meter resolution. Biomass is measured in Megagrams (1000 kg) of Carbon per hectare. Second, I also employ a lower-resolution biomass dataset which imputes biomass using the long time series of satellite images (such as greenness or other vegetation indices). There estimates yield a 5-year panel dating back to 1950 at a coarser 1 degree resolution.[4]

Absent further correction, the 1 degree resolution has significant aggregation bias. Comparing biomass in 2010 with biomass in 1980 among the subset of land which had no recorded forest cover change, 1 ton of biomass was 0.1 tons in 1980. Thus, I downscale the low resolution data to the resolution of the richer 2010 biomass vintage. Downscaling aims to match the average biomass in megatons on land with no deforestation across these two periods. That is, the downscaling factor is constant everywhere:

$$\text{Biomass}_{1980}^{\text{Downscaled}} = \frac{\sum_{\omega:\mu_{2016}^{F}(\omega)=\mu_{1980}^{F}(\omega)} \text{Biomass}_{2016}(\omega)}{\sum_{\omega:\mu_{2016}^{F}(\omega)=\mu_{1980}^{F}(\omega)} \text{Biomass}_{1980}(\omega)} \times \text{Biomass}_{1980}(\omega)$$

This downscaling likely overstates biomass in 1980 because primary forest does still sequester nontrivial levels of carbon over a 36-year period. Then, the aboveground biomass in 2010 on land with no deforestation should not be precisely equal to its biomass in 1980.

I describe my method for converting this biomass data into a measure of emissions in the main text. I provide some validation here by calculating emissions from long-run deforestation between 1982-2016 using my emissions formula. I take "remaining biomass" to be the biomass which was not converted to emissions, e.g.:

$$\text{remaining } \hat{\text{biomass}} \text{ in } 2016 = \frac{\text{forest in } 2016}{\text{forest in } 1982} \times \text{biomass in } 1980$$

4. The data combines modern, high-resolution biomass measurement with LiDAR technology with a machine learning method to back-cast biomass.

I have two cross-sections of biomass data, one in 1980 and the other in 2010: while these imperfectly match the window from the land use data, they come quite close. Thus, I can compare my predicted remaining biomass in 2016 with actual biomass in 2010. The results are given as a scatterplot in Appendix Figure A.5. When I regress true data on my model's prediction, I obtain an $R^2 = 0.55$ with a coefficient of 0.80 on the data. In aggregate, the predicted remaining emissions are 95% of actual biomass. Thus, I slightly underpredict biomass. This underprediction comes from several sources:

1. Aggregation error: deforestation is likely selecting on biomass even within the 10 km grid cell resolution, so that my remaining biomass measure is likely understating true remaining biomass.

2. Dynamics: deforestation could have occurred anywhere between 1982-2016, meaning that there is time for other secondary regrowth of shrub or cropland to take place, leading to an understatement of remaining deforestation.

3. Measurement error: there is a spatial resolution difference across the two biomass maps being compared, and downscaling the lower-resolution raster introduces some variance.

Importantly, land with no standing forest can still have non-zero biomass in the data. Then, while my approach here classifies biomass loss as emissions, it may not be deforestation. A common practice is to define forested biomass as land where tree cover exceeds a cutoff (e.g., 30 or 50% in Song et al. 2018). Quantitatively, I test sensitivity of the model to an alternate emissions function which accounts for these cutoffs: emissions are 0 if land does not meet a forest cutoff in the initial period, regardless of land use change. I do not find this matters for final model outcomes.

C.3.2 Agricultural productivity.

Agricultural productivity is measured using the FAO GAEZ maps, which have a 9-by-9 km resolution. Used extensively in prior economic research, this dataset provides information on the biophysical suitability of different crops and livestock in different regions of the world, as well as information on climate, soil, and other environmental factors (Adamopoulos and Restuccia 2022; Farrokhi and Pellegrina 2023; Costinot, Donaldson, and Smith 2016) The GAEZ reports potential yields for nearly 30 crops at two input (high- and low-input) and water availability (irrigated and rainfed) levels. The data uses climatological models to create 30-year epochs of agricultural yields. In the time period of my land use data, 1980-2016, this 30-year timescale means I effectively use a single cross-section of GAEZ data. Data for

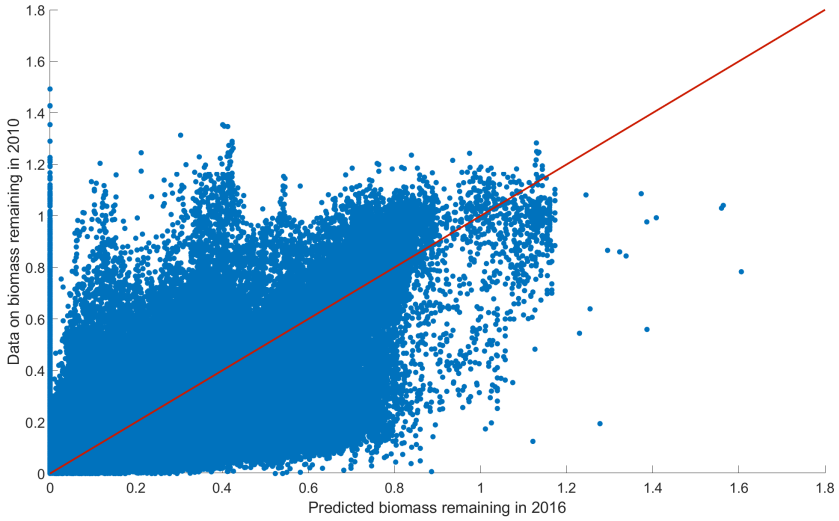


Figure A.5: Validating the empirical emissions function by comparing predicted biomass in 2016 and actual biomass in 2010.

NOTES: Each point indicates a 10 km grid cell which lost any vegetation cover between 1982-2016. The red line indicates the 45 degree line. Biomass measures are in megatons.

the main results of this paper come from high-input, rainfed yields for the climate epoch 1980-2010.

I use a single-dimensional "crop index" rather than the 30-dimensional vector of potential crop yields. Prior work often uses a weighted sum of a subset of crop yields (with weights corresponding to calorie shares of a global diet from each crop in Costinot, Donaldson, and Smith (2016), or shares of regional yields in Scott (2013)). Instead, I use a principal components analysis to reduce this large amount of potential yield data into the potential yield of a single "crop index". I compare results of the two approaches and find that they are largely comparable, but I prefer PCA for two main reasons. First, the weights used in a calorie-based or output-based weighted sum will themselves be a function of long-run relative prices and preferences, and thus introduce a source of endogeneity. Second, the principal components analysis has a convenient interpretation as maximizing the variance explained in the data in a single dimension of the data.

The first principal component, the measure I use in this paper, captures 55% of the variation in the raw data (the second component captures 24%). According to Appendix Table A.3, the procedure places greatest weight on high-calorie, staple carbohydrates. The exception is wheat, which is down-weighted significantly, perhaps because wheat is less readily grown outside of the US and the EU relative to the other staples.[5]

5. A common set of crops used in Costinot, Donaldson, and Smith (2016) also includes banana yields.

Table A.3: Loadings of principal components in the final measure.

Crop	Loading in PC1	Std. Error	Calories per kg
Sweet potato	0.45	0.001	963
Sugarcane	0.47	0.002	331
Rice	0.47	0.003	2431
Maize	0.44	0.004	2912
Oilpalm	0.30	0.006	8840
Wheat	0.29	0.010	2748
White potato	0.26	0.010	630

NOTES: Indicates the loadings of the principal components analysis onto seven staple crops as defined by the Food and Agriculture Organization. Loadings are bootstrapped over samples of 1000 5 km $\times$ 5 km grid cells. The loading represents the weight placed on the yield (in kg of dry weight) of each crop in the final principal components analysis. For example, one kilogram of sweet potatoes contributes approximately 2 times as many final crop-equivalent kilograms as a kilogram white potatoes.

I also report results for an alternate crop yield measure. Using the same input data (high-input, rainfed yields), I calculate a calorie-weighted sum of crops. Calorie shares of crops in each country comes from the FAO's Food Balance Sheet data, which provides a measure of total food supply in kilocalories and quantities for each country. I take a ratio of these values to get calories per kilogram at the country level. Oilpalm has no calorie value assigned, so I take this to be 884/100g based on the USDA Table of Weights and Measures. There is variance in calories attributed to each crop (for example, the standard deviation of 51 calories per kg bananas is 7.5% of the mean calories for bananas), which I attribute entirely to measurement error. Calories per kilogram of food are reported in Appendix Table A.3. I find that the PCA upweights sweet potatoes and sugarcane relative to calorie weighting, while underweighting oilpalm and wheat.

C.3.3 Details regarding biodiversity measurement

Biodiversity is a notoriously difficult object to quantify. My measure derives from the idea of rarity-weighted species richness, abbreviated RSR, which can be seen as an index not unlike potential yields discussed below. RSR is calculated in two steps for a given species. First, globally, one carves out suitable habitat: at minimum, this includes elevations at which the species might live. Suitability can be more detailed, depending on knowledge of the species: many birds in the BirdLife international database have a range pinned down

However, banana yields are quite spatially variable (concentrated in the tropics) and thus become upweighted by the resulting principal components analysis, leading to unrealistically high relative productivity in the tropics.

by vegetation and climate characteristics. The International Union for the Conservation of Nature, or IUCN, also maintains a detailed red list which provides such metadata on a host of threatened or endangered species. Second, for the given species' measured range, one calculates ranged species rarity by determining what fraction of that species lives in a given pixel. For example, if a bird species lives in 1000 potential tiles at a resolution of 30-by-30 meters, then the range rarity is 1/1000. My measure is an aggregate measure across the IUCN Red List and BirdLife databases, published by the IUCN. It was published in the year 2017 and sums across all species data using the 2013 vintage of the ESACCI land use data I discussed prior.

Mathematically, one can represent my range rarity measure as follows. For a given species $n \in \{1, \ldots, R\}$, each plot ω has an underlying state of suitable or un-suitable. Call this state $s_n(\omega) \in \{S, U\}$.[6] Then, I observe in my data the following metric:

$$RSR(\omega) = \sum_{n=1}^{R} \left(\sum_{\Omega} \mathbb{1}[s_n(\omega') = S | s_n(\omega) = S] dF(\omega') \right)^{-1} \tag{A.3}$$

where $R_n(\omega)$ is a species count of the species n in ω. Notably, this aggregation is species-agnostic.[7]

RSR is used often by conservation planners. It forms a useful metric which approximates the solution to an optimal planning problem whose objective is to minimize conserved area, subject to the protection of a certain amount of species. This is the "optimal site" problem. In this way, RSR is close to the shadow value of a conservation planner's cost-minimization problem of assigning a conservation state $\kappa(\omega) = 1$ as:

$$\begin{aligned} &\min_{\kappa} \sum_{\omega=1}^{N} \kappa(\omega) \\ &\quad \text{s.t.} \sum_{\omega=1}^{N} \kappa(\omega) \mathbb{1}[s_n(\omega) = S] = Q \end{aligned}$$

Subsequent steps in the conservation planning process introduce considerations such as com-

6. I treat this state as fixed for the time being, but in truth land use actions nearby will affect $s_n(\omega)$. I am limited by a single cross-section of global range data.

7. A unit of increase in the range $R_n = \sum_{\Omega} \mathbb{1}[s_n(\omega) = S] dF(\omega)$ of a single species n is identical to that of any other species:

$$\frac{\partial RSR(\omega)}{\partial R_n} = -\left(\frac{1}{\sum_{n=1}^{R} R_n} \right)^2$$

This assumption is perhaps undesirable if, for example, a threatened species is more valuable than an un-threatened species.

plementarities across nearby land, creating a difficult integer programming problem. For this reason, RSR has emerged as a popular metric. However, RSR is often computed with site-specific studies and very little global data exists outside of the IUCN categorization.

Without explicit preferences over species and a measure of costs per land, the above optimal site problem is explicitly non-economic in nature. However, it lends interpretation to a fairly singular biodiversity datapoint with global breadth: a unit of ranged species rarity can be seen as a noisy measure of the counterfactual value of land as a protected area solely set aside for biodiversity purposes. The true value would also price these marginal gains in rarity, which could in turn vary across species, but absent sufficient species data to construct these comparisons, I begin with the RSR metric.

Plots with higher values of rarity house more rare species on average. A change in $RSR(\omega)$ can be induced by either a change in the total range of the species housed on ω or the presence of the species on ω itself, creating spillovers across space. The farmer on ω can change their land use, or farmers elsewhere can do so and create spillovers onto $RSR(\omega)$. Thus, internalizing RSR entails a public goods problem, much like the emissions from deforestation to aboveground biomass – the total consumption value of aboveground biomass is a public good, not endemic to the amount of biomass on ω's own land. I leverage this public goods problem to consider the cost of biodiversity in a framework which is fairly similar to my core Pigouvian taxation framework.

C.4 Prices

My estimation procedure requires data on agricultural prices. Because my agricultural productivities and the model consider a single aggregate crop output, I primarily rely on the Food and Agriculture Organization's producer price index (PPI), an inflation-adjusted index which captures heterogeneity in farm-gate prices across countries, and within countries but across time. The underlying data used to construct this measure come from annual prices from 1991 to 2023 for 160 countries and for 262 products (though this is unbalanced at the country-crop level). The producer price index is then the classic Laspeyres price index of purchases relative to a base period of 2014-2016, when the index is set to 100:

$$p_{it} = 100 \times \sum_k \frac{p^k_{it} q^k_{i,2014-2016}}{p^k_{i,2014-2016} q^k_{i,2014-2016}}$$

where the summation index k corresponds to the set of 262 products available in the data, p^k indicate the farm-gate prices of those products, and q^k indicate production quantities.

C.5 Production and trade flows

For the estimation of the demand model and trade barriers, I use data on international production quantities and trade flows from the International Trade and Production Database for Estimation, or ITPD-E (Borchert et al. 2021; 2022). The ITPD-E is a product of the United States Trade Commission assembled to allow for explicit estimation of gravity-type demand models. It includes flows in values of commodities at the Harmonized System (HS)-2 level for 1991-2021 at an annual frequency, between 265 countries. This amounts to 170 industries grouped into four broad sectors: Agriculture is used as the A sector for this paper. I omit the mining and forestry broad sectors (HS2 codes 27-35), so that there is a mapping from my manufacturing sector to the same in the ITPD-E. A key feature of this data, as compared to comparable trade datasets like CEPII BACI used in prior agricultural macroeconomics and trade research, is it limits the use of interpolation or other statistical techniques. It also includes production data, which facilitates estimating domestic absorption.[8]

C.6 Other data sources.

There are several other data sources which are used throughout the paper. Appendix Table A.21 summarizes key raster and geospatial datasets, provides the spatial resolution of these data, and gives a brief description of each. For climate and weather controls, I construct binned heating and cooling degree day measures from the BEST satellite data. These controls are effectively 1×1 degree fixed effects as they are not at the same resolution as other data, so I do not include them in the main text estimates; they are only used in the robustness checks. I control for ERA-5 measures of precipitation using a third-order b-spline. Unless otherwise noted, I use travel time to cities from Nelson et al. (2019) as the key transportation cost control.

I additionally collect information on carbon pricing initiatives from the Global Carbon Pricing Dashboard posted by the World Bank. Commodity prices, for use in later instrumental variables strategies, come from two sources: (1) the World Bank's Pink Sheet, which provides global annual series of commodity prices for a wide range of commodities including fertilizers and various oil varieties (such as West Texas Intermediate and Brent Crude), (2) the International Fertilizer Organization's data on fertilizer usage which is a country-year panel on the use of three key fertilizer chemicals – nitrogen, phosphorus, and potassium – by country. The latter is proprietary data granted under an academic usage agreement.

8. In the agricultural sector, 57% of the 256 ITPD-E countries have at least one HS-2 code level industry's domestic production attached. In manufacturing this number drops to 32%.

I use a variety of data from the Food and Agricultural Organization (FAO). These data all are country-by-year panels for 1990-2021. Data include information on value added per worker, value add shares of labor in agriculture and forestry activities (which I use directly as a measure of agricultural value add absent further forestry-specific data), gross production in thousands of dollars USD, pesticide use in kilograms per hectare, estimated emissions of carbon dioxide in kilograms per hectare, fertilizer usage for each of three chemicals (Urea/nitrogen, phosphates/phosphorus, and potassium/potash) in kilograms per hectare, land use and landcover shares including information on irrigation usage, and finally a suite of food security indicators. All data are publicly accessible via the FAOSTAT portal.

D Land use estimation appendix

D.1 Ordinary least squares results

Table A.4 provides an ordinary least squares estimate of Equation (11). In the next three columns, I present estimates using a country-level trend to control for wage endogeneity and demonstrate my travel time instruments. The results of adding a country-level trend align with my identification argument. Because wages and other unobservable factors clear locally (within-country), adding a country fixed effect removes spurious correlation between observables η^A and τ and these local counfounders. Because input prices will generally be higher in low-productivity and high-transportation cost areas, the country fixed effect removes an attenuating bias on the estimate of γ and a bias which makes transportation costs too expensive γ_τ.

The straight-line distance instrument in table A.4 reduces attenuation of the transportation cost. Because transportation costs are correlated with the unobservable local productivity $\xi_t(\omega)$ (such as local amenities), the direction of the change in coefficients is consistent with low transportation cost areas having higher unobservable non-agricultural productivity. However, without further remedies for correlation between $\eta^A(\omega)$ and even p, the residual variation in $\tau(\omega)$ now loads onto the estimate of γ, changing its sign.

D.2 First-stage estimations for IV

Here, I provide the first stage of agricultural prices against my WTO accession-based variety instrument. Theoretically, I argue that greater exposure to the WTO accession shock should result in a larger *kink* in prices. I focus on a kink in prices because levels of prices differ for many unobservable regions – e.g., wage differences. A kinked first stage is also consistent with the second-stage regression, which is estimated with a country fixed effect.

Appendix Table A.5 illustrates the first stage relationship between the demand shifter and farm-gate prices, estimating the following equation.

$$p_{it^*} = \beta V_i + \lambda_{j^*(i)} + \epsilon_{it^*}$$

In this static first stage, I construct a mixed cross-section so that the year for country i corresponds to the year in which its variety-matched partner $j^*(i)$ accessed the WTO (event-time 0, given by calendar year t^*). For example, Brazil is matched with Ecuador, which accessed the WTO in 1996, so the mixed cross-section includes Brazil in this year.

The average variety match in the data is around 0.91, indicating 75 manufacturing sectors

Table A.4: Naïve OLS regression of (11).

Coefficient on...	(1) OLS	(2) Travel Time Instrument	(3)	(4)
γ, revenues	0.0494***	-0.3874***	0.6027***	0.9083***
	(0.0127)	(0.0143)	(0.0162)	(0.0253)
$\gamma\gamma_\tau$, transportation costs	-0.0374***	-0.0586***	-0.0317***	-0.0290***
	(0.0003)	(0.0004)	(0.0005)	(0.0005)
$\gamma\phi^{FA}$, lagged forest shares	-7.729***	-7.811***	-7.358***	-7.370***
	(0.0127)	(0.0129)	(0.0144)	(0.0145)
R^2	0.42977	0.42406	0.46904	0.47074
Observations	9,855,517	9,233,819	9,233,819	9,233,819
Dependent variable mean	-0.56672		-0.56672	-0.56672
Wald (1st stage), γ_τ		156,128.5	219,784.2	220,381.9
F-test (1st stage), γ_τ		8,355,393.4	7,428,147.5	7,000,292.8
country fixed effects			✓	
country-level trends				✓

NOTES: Regressions implement 11. Panel land use data runs from 1995 to 2015 in the ESACCI data: I discard the last 5 years of the data due to a change in data reporting. Standard errors are clustered at the 50 km block level. Columns (2)-(4) reflect two-stage least squares estimates, where travel time (in 100 minutes) is instrumented by the straight-line distance to a city (in kilometers).

demanded by the best-matched WTO accessing country are produced in the average paired non-accessing country. Then, the estimated first stage coefficient suggests that an additional sector in common produces a \$0.65 increase in agricultural prices.

To ensure effects are not uniquely driven by any one accession, I re-estimate this first-stage model by sequentially dropping individual accessing countries from the estimation. In particular, if only a subset of accessions are truly unanticipated, this procedure should result in major changes in the coefficient of interest when I drop one of these cases. I find that the most important accession by far, in terms of its impact on downstream prices, is the Russian accession. Ignoring the Russian accession reduces the coefficient magnitude by 33%, implying that countries which were relatively exposed to Russia experienced particularly large gains.

I additionally run a generalized event study model of prices. The event study is critical to assessing the two identifying assumptions, that (1) there is no anticipation of the policy, which could indicate that there are unobservable supply shocks prior to treatment driving treatment effects ex post, and (2) trends prior to policy implementation are approximately parallel, indicating that they would have likely continued to be parallel after treatment. I run the following event study regression using the two-stage staggered event study framework from (Gardner et al. 2024):

$$\Delta p_{it} = \sum_{\tau=1995-g_{j^*(i)}}^{2019-g_{j^*(i)}} \beta^{\tau} V_i \times \mathbb{1}[t - g_{j^*(i)} = \tau] + \lambda_{j^*(i)} + \epsilon_{it}$$

This places the shock in an event study-style framework around the date of accession of the partner country $g_{j^*(i)}$. The event study the hypothesized direction of the first stage: there is a clear kink in prices immediately in the accession year, with a statistically positive effect persisting up to 10 years out. Further, pre-treatment effects show no clear trend nor sign of anticipation.

I further provide evidence that my hypothesized mechanism is at play. Using a similar event study specification, I now explore the impact of accession variety on the volume of trade. I estimate:

$$\Delta E^h_{ijt} = \sum_{\tau=1995-g_{j^*(i)}}^{2019-g_{j^*(i)}} \beta^{\tau} V_i \times \mathbb{1}[t - g_{j^*(i)} = \tau] + \lambda_{j^*(i)} + \epsilon_{it}$$

In countries with relatively higher variety match, trade flows should increase as a conse-

Table A.5: First stage of the variety-matched WTO accession instrument.

Farm-gate producer price index	
WTO Variety Index	46.44**
	(22.46)
R^2	0.73257
Observations	133
F-statistic	20.088
Accessing partner fixed effects	✓

NOTES: First stage regression of the FAO producer price index measure (with 100 = the index value in 2015) against the best-case variety match with a country that accessed the WTO. The data year corresponds to the year in which a given partner accessed the WTO (year 0 in event-time, which can vary across units). There are 204 countries in the sample with 22 potential accessing partners. 5 countries which accessed the WTO – Kazakhstan, Liberia, Tonga, Yemen, and the Seychelles – only have 1 variety-matched partner and are thus perfectly predicted by accessing-partner-level fixed effects.

quence of higher exposure to accession under a demand-side stor. Figure A.7 confirms that trade flows rise as a consequence of accession, again without a clear pre-trend nor sign of anticipation. With prices kinked upwards and trade flows rising after accession, I validate that the accession primarily operates as a shock to aggregate demand.

D.3 Spatial first differences estimation

I report the results of the spatial non-linear least squares estimation, based on (12), in Table A.6. Top row estimates are non-linear least squares estimates, while bottom row estimates introduce the variety shifter instrument. Differences in the X-difference correspond to differences between neighbors no more than 15 kilometers apart (recall grid cells are 10 km apart) and sharing the same latitude. Similarly, Y-differences compare neighbors in the same longitude. Finally, diagonal differences compares the grid cell contiguous neighbor to the top right of the target grid cell. AR(1) persistence coefficients correspond to the average estimated correlation between deforestation outcomes of neighbors.

After a spatial first difference, concerns about modeling landowners are arguably mitigated. Scott (2013) points out that myopic models, where landowners cannot respond to the future value of their land, return biased estimates of a land use elasticity because they neglect the option value of future land use change. In particular, the Euler Conditional Choice Probability methods in this vein have a left-hand side which depends both on the probability of deforestation in the current period and forest regrowth in the next period. This second component is missing from my myopic specification. However, after spatial first

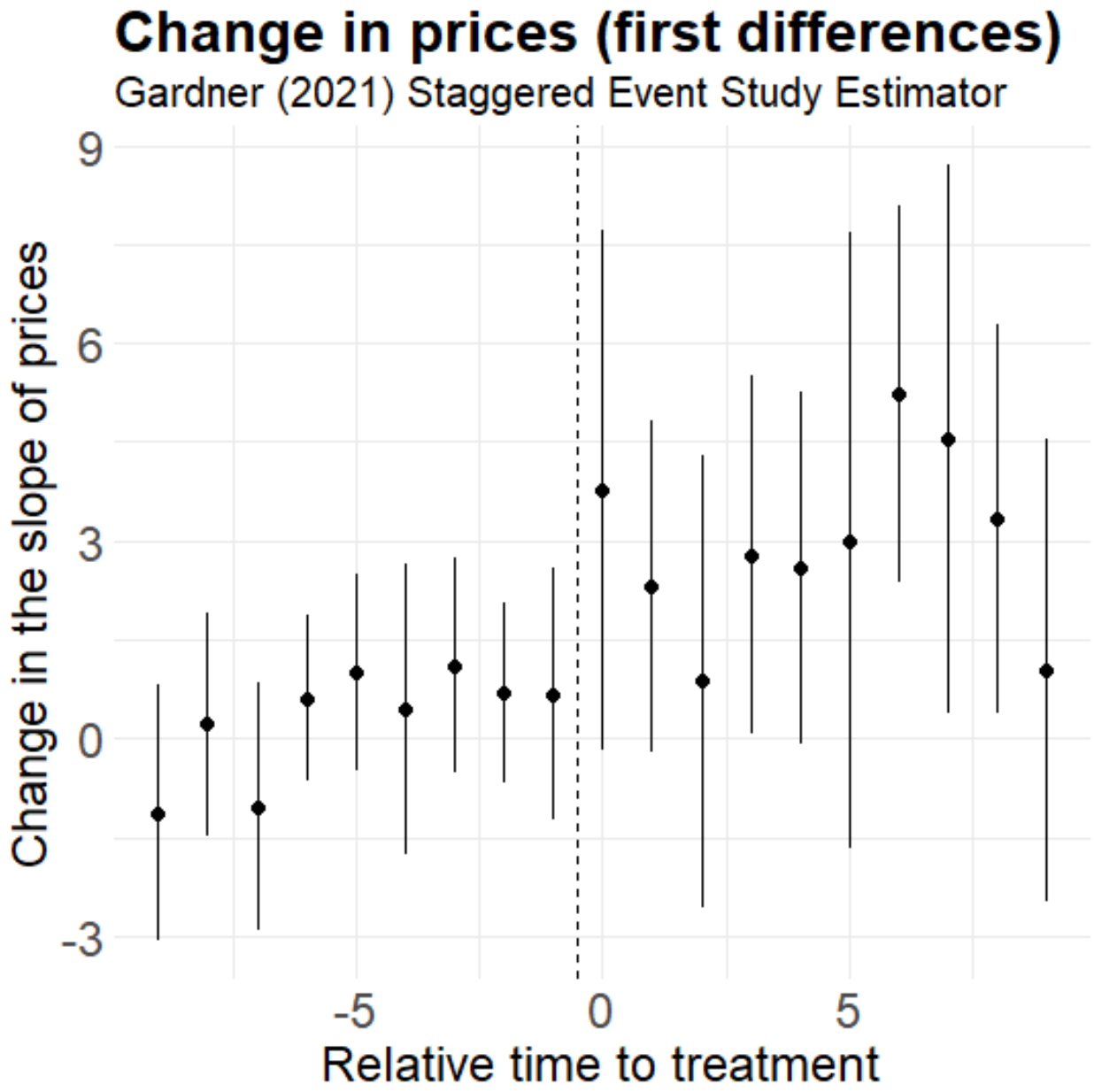


Figure A.6: Event study estimates of the impact of a high-variety match WTO accession on local farm-gate prices.

NOTES: First stage regression of the first-difference of FAO producer price indices measure (with 100 = the index value in 2015) against the best-case variety match with a country that accessed the WTO. There are 204 countries in the sample with 22 potential accessing partners.

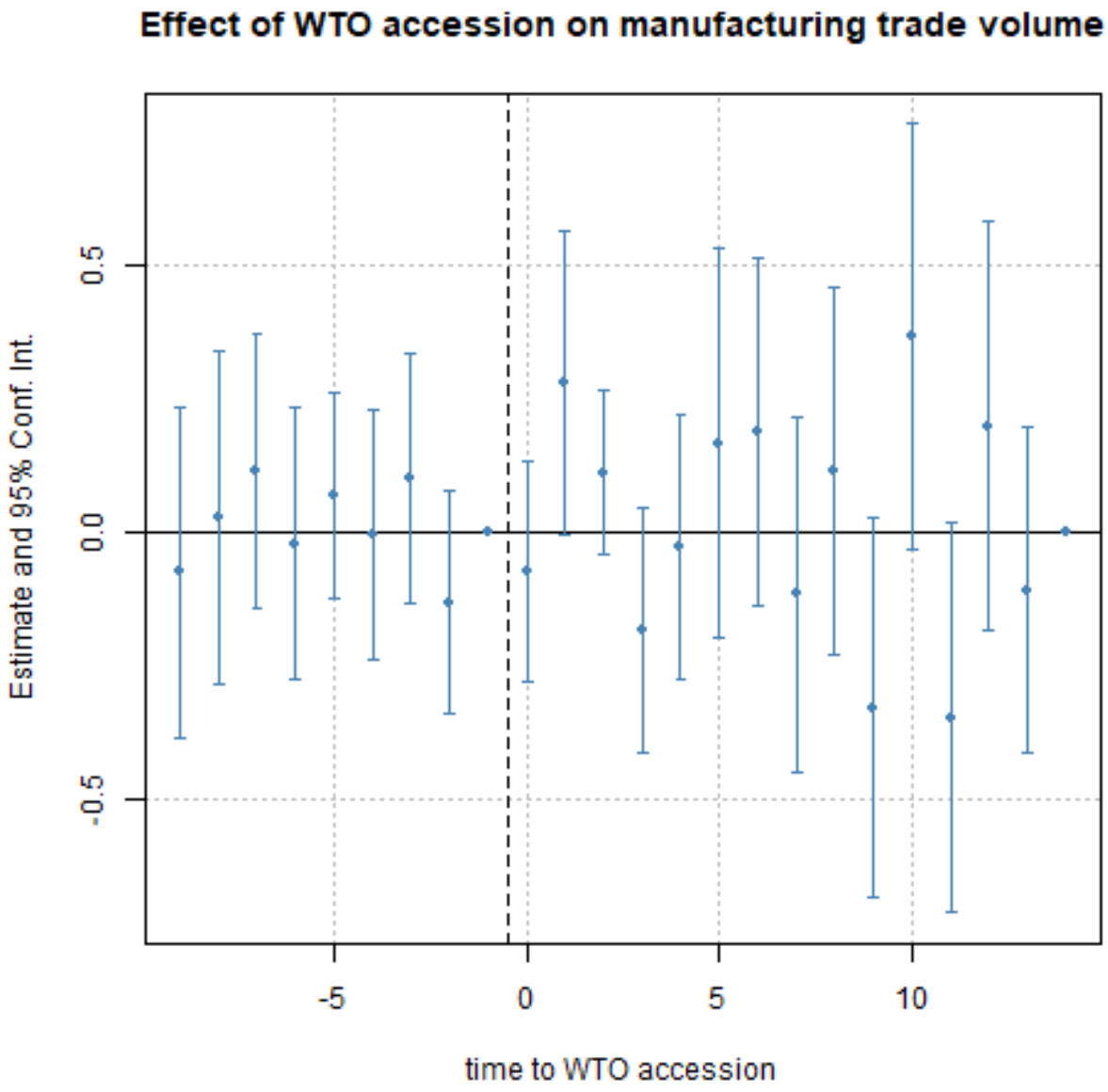


Figure A.7: Event study estimates of the impact of a high-variety match WTO accession on trade flows.

Notes: First stage regression of levels of trade flows against the best-case variety match with a country that accessed the WTO. There are 204 countries in the sample with 22 potential accessing partners.

differencing, regressions which include the dynamic correction term no longer differ from regressions which omit this term.

Table A.6: Summarizing spatial first differences estimates of landowner's problem, equation (12).

Outcome: Relative agricultural land use share			
	Diagonal	Y	X
γ, Revenue	0.77 (0.00, 2.50)	1.06 (0.17, 3.60)	1.17 (0.29, 3.99)
Transportation cost per minute	−0.04 (−21.48, 0.00)	−0.04 (−0.97, −0.01)	0.01 (0.01, 0.58)
Cost of deforestation	−9.48 (−79.80, −3.11)	−7.32 (−59.31, −2.66)	−4.53 (−15.05, −1.26)
AR(1) Persistence	0.46 (0.22, 1.12)	0.20 (0.08, 0.23)	0.38 (0.21, 0.81)
	X, IV	Y, IV	Alt. Func. Form.
γ, Revenue	1.05 (0.53, 2.40)	1.04 (−0.03, 4.30)	1.84 (0.39, 5.05)
Transportation cost per minute	−0.01 (−0.06, 0.52)	−0.04 (−1.80, 0.00)	−0.21 (−5.95, 7.42)
Cost of deforestation	−6.10 (−11.18, 1.79)	−7.38 (−191.48, −1.76)	−3.82 (−19.97, −1.41)
AR(1) Persistence	0.39 (0.10, 0.71)	0.39 (0.06, 0.77)	0.36 (0.19, 0.71)

NOTES: Regressions implement 12. Panel land use data runs from 1995 to 2015 in the ESACCI data: I discard the last 5 years of the data due to a change in data reporting. Standard errors are computed with 100 iterations of a bootstrap of 1000 50×50 km blocked clusters sampled with replacement with robust confidence intervals from Efron (1987). Column headers for (1)-(4) indicate the direction of spatial differencing. Column (4) removes country-by-year fixed effects from the specification. Column (5) uses the IV GMM criterion with two instruments: straight-line distance to market and a variety-weighted WTO accession instrument. Column (6) uses an alternative functional form for returns.

D.4 Robustness: alternate instrument

I propose an alternate instrument which leverages distinct variation from the variety shifter discussed in Appendix D.2. I rely on data from the World Bank's Distortion to Agricultural Incentives data. This data reports the sum total of output price distortions – e.g., input subsidies, border adjustments, and land taxes – as a "nominal rate of assistance" on top of the farm-gate price received by a farmer. The data is discussed extensively by Adamopoulos and Restuccia (2022). A drawback of this data, however, is its recency: the dataset was discontinued in 2007.

I use the general nominal rate of assistance across domestic and border policies ('nra_covt' in the data) as an instrument for prices. Recent work has demonstrated that future changes in policy can be predicted by domestic yield shocks, so I use a simple Hausman instrument which averages others' nominal rates of assistance in the same year. This instrument is exogenous if other countries do not respond on average to supply shocks in a given country (which would be invalidated if there are large-scale subsidy races, where a change in one's own subsidy is simultaneous with that of many other countries).

$$Z_{it}^{DAI} = \frac{1}{J-1} \sum_{j \neq i, j \in \{1,\ldots,J\}} NRA_{it}$$

The domestic nominal rate of assistance has a clear theoretical relationship with prices in perfectly competitive agriculture. Within a country, raising the average rate of assistance will raise farm-gate receipts, driving up supply, and lowering prices. This channel can be hidden in a cross-country regression when countries have different price indices for other reasons – like differing agricultural yields – so I show my first stage with and without country fixed effects for clarity.

Results are shown in A.7. As expected, the Hausman-style distortions have a negative impact on prices after controlling for country fixed effects. The sign of the mechanism is robust to first differences as well. In general, the rate-of-assistance instrument is reasonable with an F-statistic of 12.5 without country fixed effects.

Using this alternate instrument results in a slightly larger estimate of γ than the one reported in the main text. Table A.8 reports the OLS (identical to Table 2) and two-stage least squares estimates with this instrument. The estimate of the deforestation elasticity is significantly larger, though it cuts the sample size in half due to the length of the panel. The benefit of the WTO instrument is the scope: demand shocks are global in scale, and the identification comes off of the dose. The benefit of the NRA instrument is its timing, which allows for higher (temporal) resolution identification.

Table A.7: First stage with alternate Hausman nominal rate of assistance instrument.

	Farm-gate producer price index				
	OLS		1-Period Lag		First Differences
	(1)	(2)	(3)	(4)	(5)
Nominal rate of assistance	1.385***	-9.410**	1.318***	-6.930*	-4.677**
	(0.3018)	(4.419)	(0.2922)	(3.864)	(1.995)
R^2	0.16561	0.71882	0.17020	0.74424	0.01535
Observations	939	939	877	877	876
F-test	12.306	79.250	12.511	88.753	0.95114
country fixed effects		✓		✓	

NOTES: Observations are country-years. Estimates the first-stage effect of a Hausman instrument of nominal rates of assistance to crops on prices. Columns (1) and (2) use the untransformed nominal rate of assistance, (3) and (4) the lagged value, and (5) regresses temporal differences in price on differences in the NRA.

Table A.8: Estimates of the landowner's problem using the nominal rate of assistance instrument.

Relative agricultural land use share (logs)		
	(1)	(2)
	OLS	IV
γ , Revenue	1.119***	1.926***
	(0.0355)	(0.0478)
$\gamma \times c(\omega)$, Travel time	-0.0233***	-0.0207***
	(0.0005)	(0.0005)
$\gamma \times \phi^{FA}$, Lagged forest share	-7.332***	-7.142***
	(0.0182)	(0.0197)
R^2	0.47784	0.45826
Observations	4,549,776	1,842,379
Dependent variable mean	-0.59186	-0.83088
F-test	256.49	432.93
Wald (1st stage), γ		57,686.6
country-level trends	✓	✓

NOTES: Estimates Equation (11). Observations are grid cell-years. Standard errors clustered at the 50 km block level. Instrument is leave-one-out domestic agricultural incentives.

D.5 Robustness: many land uses

Another threat to identification comes from model misspecification. In particular, the model I focus on presents a world with only two land uses, when in reality, deforestation can contend with returns to pasture and urban areas as well. However, satellites are notoriously poor at discerning grassland, cropland, and forest at scale, with most high-fidelity datasets either limiting their attention to a forest-or-nonforest classification task or limiting the area over which their classifier operates (e.g., just the Brazilian Amazon). Consequently, I must rely on aggregate data to get the best possible estimates of land use outside of forest and agriculture.

Aggregate country-by-year level land use data comes from the FAO. I calculate the share of pasture in a given country as the difference between its agricultural land use share and its cropland land use share. I then take the difference between the total land use share (1) and the sum of crop, pasture, and forested land use to get an "other" land use share, which includes both bare areas (tundra and desert, for example) and urban, settled regions.

I leverage country-level land use data in a multinominal logit model. Country level data is denoted with a bar, e.g. country-level average potential yields are $\bar{\eta}_i^A$. Agricultural returns given the vector of initial states s_{i0} now depend on the *average* return in country i, $\pi_{it}^A(s_{i0})$ across individual plots ω in country i. Pasture and Other land uses each have a different switching cost than agriculture, indexed by coefficients ϕ^h on lagged forested land use shares $\bar{f}_{i,t-1}$.

$$\bar{\pi}_{it}^h(s_{i0}) = \bar{\epsilon}_{it}^h + \begin{cases} 0 & \text{if } h = F \\ \gamma[p_{it}\bar{\eta}_i^A - \gamma_\tau \bar{c}_i + \phi^{FA}\bar{f}_{i,t-1}] + \xi_{it}^A & \text{if } h = A \\ \gamma\phi^P \bar{f}_{i,t-1} + \xi_{it}^P & \text{if } h = P \\ \gamma\phi^O \bar{f}_{i,t-1} + g(\text{Population}_{it}) + \xi_{it}^O & \text{if } h = O \end{cases}$$

Each land use is associated with a quality shifter at the country-year level and a modeled return, ξ_{it}^h for land use h. To estimate this aggregate land use model, I begin with a simple linear model. Define the left-hand-side variable of the model Y_{it}^{hk} by the average profit in the agricultural sector, e.g.,

$$Y_{it}^{hk} = \log(\bar{\mu}_{it}^h) - \log(\bar{\mu}_{it}^k)$$

The right-hand side of the model uses country-by-year averages of the variables measured at the plot-level. Importantly, this model thus throws away all subnational variation. I include

Table A.9: Results of estimation with expanded multinomial choice problem.

Multinomial logit, Y_{it}^{hk}	
Revenue coefficient	1.6^{***}
	(0.12)
R^2	0.17079
Observations	13,896

NOTES: Multinomial logit estimator for the land use problem on a country-by-year dataset. Standard errors clustered at the country level.

country fixed effects to account for level differences in prices. I include year fixed effects to be consistent with the myopic model, estimating purely off of within-year variation. I obtain a stacked regression of the form:

$$Y_{it}^{hk} = \bar{\pi}_{it}^{h} - \bar{\pi}_{it}^{k} + \lambda_i + \lambda_t + \epsilon_{it}^{hk}$$

Standard errors are clustered at the country level to account for serial correlation in the panel. The resulting multinomial logit revenue elasticity in Appendix Table A.9 is extremely similar to the estimate from the main text.

D.6 Robustness: alternate agricultural yields

I re-run the regression from Equation (11) using calorie-weighted yields (see the Data Appendix, Section C.3.2 for details). I test both high and low input yield assumptions from the FAO. Results are broadly consistent with the preferred specification in Table 2, as demonstrated by the elasticity equivalent. Deforestation responds more to high input yields on average.

Relative agricultural land use share (logs)		
	(1)	(2)
	OLS	IV
Low-input Yields		
γ, revenue	0.262***	0.367***
	(0.008)	(0.009)
$\gamma \times c(\omega)$, travel time	-0.0233***	-0.0219***
	(0.0005)	(0.0005)
$\gamma \times \phi^{FA}$, lagged forest share	-7.075***	-7.036***
	(0.0190)	(0.0189)
R^2	0.47380	0.47356
Wald (1st stage), γ		77,199.4
Avg. elasticity equivalent	0.04	0.05
High-input Yields		
γ, revenue	0.15***	0.30***
	(0.009)	(0.010)
$\gamma \times c(\omega)$, travel times	-0.0252***	-0.0239***
	(0.0005)	(0.0005)
$\gamma \times \phi^{FA}$, lagged forest share	-7.162***	-7.152***
	(0.0188)	(0.0187)
R^2	0.47277	0.47233
Wald (1st stage), γ		91,558.8
Observations	1,112,359	1,112,359
country fixed effects	✓	✓
Avg. elasticity equivalent	0.04	0.08

Table A.10: Calorie-weighted yields.

NOTES: Re-estimates (11) with alternate yield measures which are sums of calorie-weighted potential yields of crops. Top panel is under low-input assumption, while bottom panel is under high-input assumptions. Column (1) is an OLS estimate while Column (2) uses the WTO accession variety instrument to instrument for prices.

D.7 Heterogeneity analysis I: Observables

In this section, I include Appendix Table A.12 which reports the results of instrumented least-squares models of the landowner's problem in equation (1). Each model perturbs the baseline specification to allow for richer heterogeneity in the parameters (γ, γ_τ, ϕ^{FA}), or to control for potential confounders. I discuss these results in the main text in Section 5.

In Appendix Table A.11, I also provide some brief descriptives regarding heterogeneity in switching costs across countries. These switching costs come from the country-specific slopes in column (1) of Appendix Table A.12, so that a negative coefficient implies a higher switching cost. Cross-country switching costs are higher in countries with less forest, suggesting that there may be incentives to have non-zero forest cover aside from carbon. Higher GDP-per-hectare countries also have lower switching costs: I attribute these differences to enforcement quality. Finally, highly productive timber producers in 1990 have lower switching costs today. In unreported regressions, I do not find significant correlations between switching costs and land protection.

In ongoing robustness checks, I use a local linear forest to construct nonparametric observable heterogeneity in γ (Wager and Athey 2018). Initial results support significant nonlinearities in observables but do not support dramatically different point estimates.

Table A.11: Summarizing cross-country heterogeneity in switching costs.

	Switching cost coefficient			
	(1)	(2)	(3)	(4)
Log forest area, thousands ha	-0.1211**	-0.1489***	-0.1542***	-0.0795
	(0.0557)	(0.0567)	(0.0548)	(0.0513)
Log population density		-0.0865		
		(0.0584)		
Log GDP per hectare			-0.0974***	
			(0.0357)	
Log woodfuel, cub. m per ha in 1990				-0.0748***
				(0.0229)
R^2	0.13705	0.15572	0.18331	0.16315
Observations	116	116	116	92
Average RHS	28,527.0	0.0017	48.85	0.0002
SD RHS	97,663.7	0.0062	391.5	0.0004

NOTES: Regresses the log of estimated switching costs ϕ^{FA}/γ in dollars per hectare on the log of a number of explanatory variables from the Food and Agriculture Organization's 2010 Forest Resources Assessment. GDP per hectare is in thousands of dollars per hectare. Population density is in thousands per hectare. Woodfuel is measured in cubic meters. Standard errors are at the country (observation) level.

Table A.12: Testing for sources of heterogeneity in Equation (1).

	Relative agricultural land share					
	(1)	(2)	(3)	(4)	(5)	(6)
γ, Revenue	1.606***	1.514***	1.494***	1.539***	1.349***	1.925***
	(0.0352)	(0.0386)	(0.0360)	(0.0360)	(0.0379)	(0.0365)
$\gamma \times c(\omega)$, Travel time	-0.0240***		-0.0209***	-0.0209***	-0.0153***	-0.0150***
	(0.0005)		(0.0005)	(0.0005)	(0.0006)	(0.0006)
$\gamma \times \phi^{FA}$, Lagged forest share		-7.286***	-7.272***	-7.283***	-7.309***	-8.057***
		(0.0185)	(0.0184)	(0.0183)	(0.0190)	(0.0267)
Travel time × potential yields			-12.13***			
			(0.6121)			
Travel time × revenues				-0.2029***	-0.1834***	-0.3054***
				(0.0106)	(0.0107)	(0.0114)
R^2	0.49124	0.48271	0.47891	0.47911	0.48938	0.49322
Observations	5,138,107	5,138,107	5,138,107	5,138,107	5,138,107	5,138,107
country fixed effects	✓	✓	✓	✓	✓	✓
accessing partner fixed effects	✓	✓	✓	✓	✓	✓
Lagged forest × country	✓					
Travel time × country		✓				
regrowth rates × country					✓	
aboveground biomass × country						✓

NOTES: Standard errors clustered at the 50×50 km block-level. Left-hand side variable is the log odds ratio of agricultural and forested land use shares. I use land use data between 1995-2015. All columns are esitmated via two-stage least squares, using the variety shifter instrument described in the main text Section 4.

D.8 Heterogeneity Analysis II: unobservables

As a final robustness check, I explore unobservable heterogeneity in the key parameters γ and ϕ^{FA}. Heterogeneity in γ is important to accurately assessing distributional consequences of a carbon tax. Further, if unobserved heterogeneity in returns is correlated with potential emissions, then ignoring heterogeneity can bias the estimate of avoided deforestation. Heterogeneity in ϕ^{FA} is valuable for understanding incidence: the perceived switching cost measure captures the pre-intervention, business-as-usual private cost of deforestation. Comparing these switching costs with policy incidence can give insights into where the policy has greatest bite.

Unobservable heterogeneity in the idiosyncratic shock γ. I adopt a semiparametric estimator for unobservable heterogeneity whereby I assume that γ is described by a distribution $\mathcal{F}(\mu^\gamma, \sigma^\gamma)$. My preferred specification will take γ to be drawn from a normal distribution. For plot-level draws from the standard normal distribution $x(\omega) \sim \mathcal{N}(0,1)$,

$$\gamma(\omega) = \mu_1^\gamma + \sigma_1^\gamma x(\omega)$$

Because I take choice probabilities directly from the data, estimation does not require the usual expectation step of an expectation-maximization algorithm.

I conduct heterogeneity analysis with respect to potential unobservable variation through the following maximum likelihood algorithm. Define country-level observed land use shares of h as s_{it}^h and the average share of country i's land in land use h as implied by Equation (1) as $\hat{\mu}_{it}^h$

1. For a specific distribution $\gamma \sim \mathcal{F}(\mu^\gamma, \sigma^\gamma)$, guess a set of nonlinear parameters $(\mu^\gamma, \sigma^\gamma)$
2. Draw for every plot ω a value $\gamma(\omega)$ according to $\mathcal{F}$
3. Recover the linear parameters (γ_τ, ϕ^{FA}) using the following residualized regression

$$\frac{Y_t(\omega)}{\gamma(\omega)} - p_{it}\eta^A(\omega) = \gamma_\tau c(\omega) + \phi^{FA} f_{t-1}(\omega) + \xi_t(\omega) + \epsilon^{FA}(\omega)$$

   I report a spatially differenced result which is robust to correlation between the lagged dependent variable f_{t-1} and unobservables $\xi_t(\omega)$.

4. Construct a GMM criterion interacting unobserved land quality with the WTO accession instrument $\mathbb{E}[V_i \xi_{it}] = 0$.
5. If the guess of nonlinear parameters results in a value of the criterion within some tolerance, conclude. Else, draw a new guess.

Results are summarized in Table A.13. I test three models: the single normal distribution I described above, a richer mixture of two normal distributions, and a single normal distribution where the unobserved shock to $\gamma(\omega)$ can be correlated with transportation costs. That is, in column (3),

$$\gamma(\omega) = \mu_1^\gamma + [\sigma_1^\gamma + \sigma_\tau^\gamma \log(c(\omega) + 1)]x(\omega)$$

This last specification captures the idea that distance is in effect a bundled treatment on agricultural returns. Not only are there physical transportation costs, but more distant land is also associated with greater downstream market power in trucking and value chains (Allen et al. 2022; Sayre 2023). Greater downstream market power means farmers are more likely to sort towards high-productivity land, as low-productivity land is less competitive with downstream markups, raising the elasticity. This reflects the old adage: the monopolist is the conservationists' best friend.

All three specifications return similar estimates of the mean revenue coefficient. Indeed, columns (1) and (2) are essentially identical, suggesting the model rejects richer heterogeneity. However, the distance-correlated specification in (3) reveals a much larger variance in the revenue coefficient. This large variance in the revenue coefficient does not impact other coefficient estimates: it is truly residual variance in the revenue-land use relationship which could not be captured by a purely linear model. A consequence of this additional heterogeneity is that land further from market, is, if anything, more elastic to a carbon price, because it is more attentive to the potential revenue of land than to idiosyncratic revenue shocks.

I perform two robustness checks. First, in column (4), I allow for spatial correlation beyond transportation costs by allowing for a covariance between my draw and my neighbor's draw.

$$\gamma(\omega) = \mu_1^\gamma + [\sigma_1^\gamma + \sigma_\tau^\gamma \log(c(\omega) + 1) + \sigma_2^\gamma x(d\omega)]x(\omega)$$

Spatial correlation-robust estimates increase the average realization of $\gamma(\omega)$ by 20%. If a neighbors' realization of ϵ is 1% lower, I see a 0.02% pass-through onto the variance of the landowners' own shock. Correlation attenuates γ because a high-yield landowner may deforest less due in part to its neighbors shock.

Second, in column (5), I check for cross-country variance in γ explained by (1) GDP per capita or (2) the labor share of value added in agriculture. Labor share of value added is a technological measure which proxies for the presence of smallholder, labor-intensive farming. Labor-intensive farming is theoretically more "exposed" (less adaptive) to shocks (Foster and Rosenzweig 2021). GDP per capita proxies for the average income of downstream

consumers and is used as a control to compare two countries with similar income but different agricultural input mixes. Thus, I adopt the specification:

$$\gamma(\omega) = \mu_1^\gamma + \mu_2^\gamma \log(\text{VA share}) + \mu_3^\gamma \log(\text{GDP per capita}) + [\sigma_1^\gamma + \sigma_\tau^\gamma \log(c(\omega)+1) + \sigma_2^\gamma \log(\text{VA share})]x(\omega)$$

Consistent with this story, I find a 1% increase in the labor share of value-added decreases γ by 0.12%.

Table A.13: Summary of unobserved heterogeneity results.

	Land use shares				
	(1)	(2)	(3)	(4)	(5)
R^2	0.25879	0.25879	0.25949	0.20364	0.26054
Observations	12,156	12,156	12,156	6,362	11,971
Average revenue coefficient	2.583	2.583	2.571	2.946	2.571
Std. dev. revenue coefficient	0.0003	0.0003	0.1377	0.1028	0.1623
10th pct. revenue coefficient	2.582	2.582	2.442	2.811	2.385
90th pct. revenue coefficient	2.583	2.583	2.725	3.057	2.771
Num. normals in mixture	1	2	1	1	1
$\gamma(\omega)$ depends on $c(\omega)$			✓	✓	✓
$\gamma(\omega)$ depends on neighbors				✓	
$\gamma(\omega)$ depends on value-add share					✓

NOTES: Estimates of nonlinear model of Equation (11) with unobservable heterogeneity in γ. Standard errors are bootstrapped over 100 draws, clustered at the 50 km × 50 km block level, with bias correction (Efron 1987).

D.9 Endogenizing production decisions

I maintain key assumption in the main text: landowners produce a fixed quantity $\eta^A(\omega)$ given by the potential yield of the plot ω in which they are located. A natural question arises: if landowners could adjust their production decisions in response to a carbon tax, how would this affect food production, emissions, and welfare in my key counterfactuals?

I do not treat a full production function for agriculture in this paper. Instead, I allow the landowner to make an additional endogenous production decision. Should the landowner choose to produce agriculture, they may then choose to produce irrigated agriculture or rainfed agriculture. This decision reflects a capital intensity decision, and gives landowners an additional margin on which to respond to a carbon tax.

In keeping with my strategy of linking landowners' decisions directly to the data, I choose this endogenous decision because I have yield data under irrigated and rainfed assumptions in the FAO GAEZ dataset.

To model this additional dimension, I introduce a nested choice problem for the landowner. In an inner nest, landowners face a choice between irrigated and rainfed farming. Irrigated farming earns a return $\pi^{HI}(\omega; s_0(\omega))$, while rainfed farming earns $\pi^{LO}(\omega; s_0(\omega))$. In the outer nest, landowners decide between agriculture and forest, just as before in Equation (1).

I now parameterize the returns in the inner nest. High-input irrigated farming yields (weakly) more output $\eta^{HI} - \eta^{LO} \geq 0$ per input of labor, but requires a fixed investment $\iota_0 + \iota_i$ upfront. For the time being, I assume irrigated firms face the same plot-level costs as non-irrigated firms, $c(\omega)$. The irrigated return is thus:

$$\pi^{HI}(\omega; s_0(\omega)) = p_i \eta^{HI}(\omega) - c(\omega) - \phi^A(\omega; s_0(\omega)) - a_i w_i - \iota_0 + \iota_i$$

whereas the rainfed return is:

$$\pi^{LO}(\omega; s_0(\omega)) = p_i \eta^{LO}(\omega) - c(\omega) - \phi^A(\omega; s_0(\omega)) - a_i w_i$$

Thus, the landowner's inner nest choice problem trades off irrigated and rainfed land uses as:

$$\pi^A(\omega; s_0(\omega)) = \max_{k \in \{HI, LO\}} \pi^k(\omega; s_0(\omega)) + \nu_i^k(\omega)$$

I assume that ν_i has a generalized Type I extreme value such that draws within the agricultural nest are correlated with factor $\lambda \in (0, 1)$, and the overall variance is γ. Then,

the within-agriculture conditional choice probability is given by a similar choice probability to that analyzed in the main text:

$$\mu^{HI|A}(\omega; s_0(\omega)) = \frac{\exp(\pi^{HI}(\omega; s_0(\omega)))^{\frac{\gamma}{\lambda}}}{\exp(\pi^{HI}(\omega; s_0(\omega)))^{\frac{\gamma}{\lambda}} + \exp(\pi^{LO}(\omega; s_0(\omega)))^{\frac{\gamma}{\lambda}}}$$

However, the outer nest trades off the inclusive value of producing either form of agriculture and the value of the natural use (forested) land use. The inclusive value of agriculture is given by:

$$\pi^{A}(\omega; s_0(\omega)) = \lambda \log[\exp(\pi^{HI}(\omega; s_0(\omega)))^{\frac{\gamma}{\lambda}} + \exp(\pi^{LO}(\omega; s_0(\omega)))^{\frac{\gamma}{\lambda}}]$$

Using the properties of a generalized extreme value distribution (Train 2009), I obtain two linked regression equations. First, the inner nest decision depends on relative profitability of the two types of farming. Because I assume costs are the same across farm types, this is:

$$\log \frac{\mu^{HI}(\omega; s_0(\omega))}{\mu^{LO}(\omega; s_0(\omega))} = \frac{\gamma}{\lambda} p_{it}(\bar{\eta}^{HI}(\omega) - \bar{\eta}^{R}_{LO}(\omega)) + \iota_i + \iota_0 + \epsilon^{IR}(\omega)$$

Next, the outer nest resembles that in the main text but for the revenue term. The revenue term is replaced with the inclusive value of the inner nest:

$$Y_t(\omega) = \lambda \overbrace{\pi^{A}(\omega; s_0(\omega))}^{\substack{\text{Inner nest} \\ \text{inclusive value}}} + \gamma \phi^{FA} f_{t-1}(\omega) + \beta X_t(\omega) + \xi(\omega) + \epsilon^{FA}_{it}$$

However, I lack a key left-hand side variable: irrigated land use at the grid cell level. I thus cannot observe the inner nest decision within agriculture. However, I observe aggregate land shares of irrigation in 219 countries via the Food and Agriculture Organization (which in turn collects these land use data from national statistics and harmonizes them, see Appendix C). Thus, I instead estimate an aggregate inner nest problem:

$$\log(\text{Irrigation share/Rainfed share})_{it} = \frac{\gamma}{\lambda} p_{it}(\bar{\eta}^{I}_{i} - \bar{\eta}^{R}_{i}) + \iota_i + \iota_0 + \epsilon^{IR}_{it}$$

For identification, I estimate the outer nest using the same the variety instrumental variables strategy from Section 4. The inner nest is instrumented using Z_{it}, the alternate instrument I discuss in Appendix D.4. Note that the identification challenge across both equations is the same: a simultaneity concern.

Table A.14: Results of nested logit estimation of the landowner's problem.

	(1) Inner nest	(2) Outer nest
Revenue coefficient, $\frac{\gamma}{\lambda}$	0.86**	
	(0.13)	
Within-nest correlation, λ		0.72***
		(0.28)
Standard-Errors	country	NID5
R^2	0.99235	0.16511
Observations	1,096	5,636,233
Elasticity	$\gamma = 0.39$	$\lambda = 0.72$

NOTES: First column is estimated on country-by-year data; second on grid cell-level data. Instrument in column (1) is variety shifter; in column (2) is leave-one-out domestic agricultural incentives.

Results of the two-step regression are reported in Appendix Table A.14. I make two observations. First, accounting for technology raises the estimate of the deforestation elasticity from 0.1 in the main text to 0.72 here. Second, the intensification elasticity is 33% higher than the extensification elasticity: agricultural firms are more responsive to yields when deciding technological inputs than on the deforestation margin.

I additionally allow for heterogeneity in the parameter λ. Economically, the unobservable shock across irrigated and non-irrigated land is likely more decoupled in areas which are more dry. I thus re-estimate the outer nest model from Equation (11), additionally interacting the revenue term with a measure of long-run soil water balance from the CHELSA satellite:

$$Y_t(\omega) = \lambda p_{it} \eta^A(\omega) + \lambda_{swb} p_{it} \eta^A(\omega) swb(\omega) + \cdots + \epsilon^{FA}(\omega)$$

If $\lambda_{swb} > 0$, then the average correlation between shocks on rainfed and irrigated land is lower when the base soil is dry. This result is indeed what I find: the baseline elasticity λ rises to 0.93 and the interaction term λ_{swb} is significantly positive at a 99% confidence level.

I also return country-level irrigation costs ι_i from the regression. The average such cost is \$220 per hectare. Countries with more smallholder farming, proxied by the labor share of value added, tend to have larger costs of irrigation installation as well. I report this in a scatterplot in A.8.

From a theoretical perspective, endogenous technology allows producers to adapt to a cost of deforestation either by exiting agriculture altogether or by intensifying their land or shifting technology. In general, counterfactual agriculture should thus be less affected by a Pigouvian tax as a consequence of adding the option to irrigate land. Aggregate agricultural

quantities will now depend not just on the extensification (e.g., deforestation) response, but also the extensification response, as:

$$\begin{aligned}\frac{dQ}{dt} &= \sum_{\omega=1}^{N} \frac{d}{dt}[q^{HI}(\omega)\mu^{HI}(\omega) + q^{LO}(\omega)\mu^{LO}(\omega)]\mu^{A}(\omega) \\ &= \underbrace{-\frac{\gamma}{\lambda} d(\omega)\tilde{s}^{TECH}(\omega)[\eta^{HI}(\omega) + \eta^{LO}(\omega)]\mu^{A}(\omega)}_{\text{intensification elasticity, new}} \underbrace{-\gamma d(\omega)[\eta^{HI}(\omega)\mu^{HI}(\omega) + \eta^{LO}(\omega)\mu^{LO}(\omega)]\tilde{s}(\omega)}_{\text{extensification elasticity (main text, Section 3.5)}}\end{aligned}$$

where $\tilde{s}^{TECH}(\omega) = \mu^{HI}(\omega; s_0(\omega))\mu^{LO}(\omega; s_0(\omega))$ is the logit elasticity of the inner nest. Define the relative elasticity of intensification as $\tilde{s}^{I}$ as the intensification elasticity above, the first term in $\frac{dQ}{dt}$, divided by the extensification elasticity, the second term:

$$\tilde{s}^{I}(\omega) \coloneqq \frac{1}{\lambda} \underbrace{\frac{\tilde{s}^{TECH}(\omega)}{\tilde{s}(\omega)}}_{\text{ratio of logit elasticities}} \overbrace{\frac{\mu^{A}(\omega; s_0(\omega))[\eta^{HI}(\omega) + \eta^{LO}(\omega)]}{q(\omega)}}^{\text{productive capacity at } \mu^{A}(\omega) \text{ vs. actual production}}$$

Intensification is generally more responsive than extensification – $\frac{1}{\lambda} > 1$ for $\lambda \in (0, 1)$ –and the last term is also always larger than 1 (the numerator dominates the denominator as long as yields under both technologies are positive). This result is somewhat mechanical, as intensification introduces a new "product" into the logit problem (Petrin 2002).

Consistent with this theoretical result, incorporating endogenous technology reduces food production losses under the Pigouvian tax. At baseline, the model predicts that 19.95% of cropland is irrigated, close to the true value of 19.6% in the FAO data for 2020. After introducing the same Pigouvian tax as in my counterfactual in Section 3.4, total agricultural production falls by 2.1%. Irrigation thus lowers the costs of environmental policy by 66% relative to the previously estimated 7% loss of global food production. To produce this increased food, irrigated area rises by 28 percentage points relative to business-as-usual. Because firms can adopt irrigation and produce more with less land, more agricultural land area (14%) is ceded to forest cover as a consequence of the tax.

However, these lower food costs do come at a cost to emissions. Because irrigation makes land which was previously unprofitable relatively more profitable, it reduces avoidable emissions by 2 percentage points (95% emissions reductions, relative to the previous 97%).

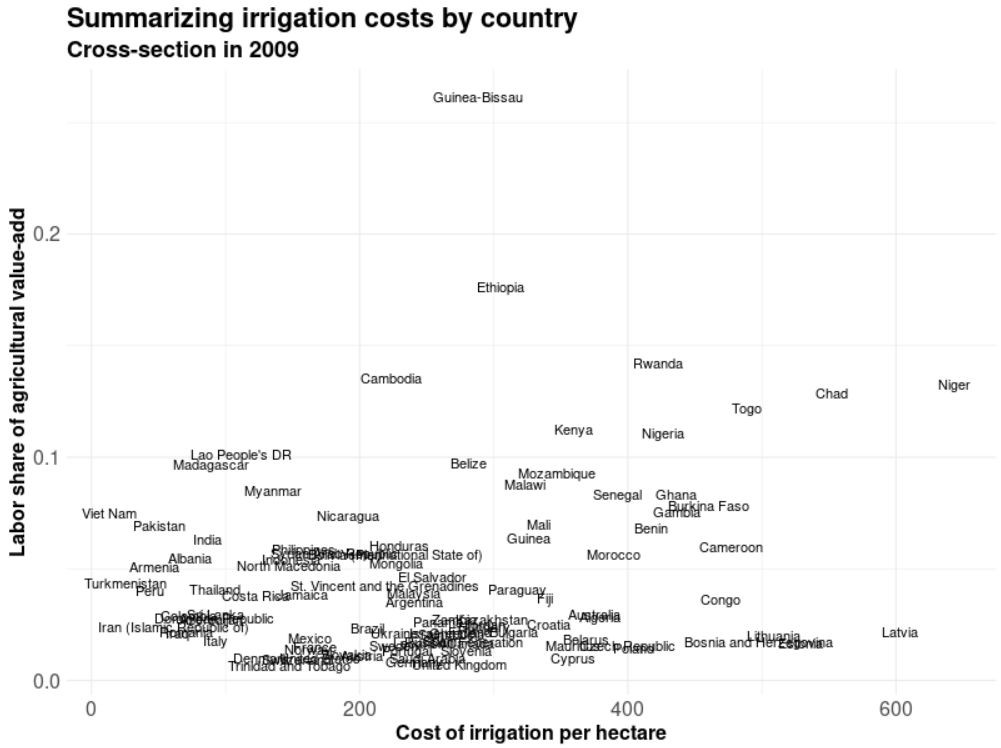


Figure A.8: Irrigation costs, ι_i, are higher in countries with a higher labor share of value added.

NOTES: Labor share of agricultural value added denoted as a fraction of total value-add based on FAO statistics.

E Additional results for demand estimation

E.1 Theory: propensity to export and import

From (14), the importer fixed effect depends on the income of the importer and a market access term:

$$\delta_j^{M,M} = (\sigma^M - \sigma)\log P_j^M + \log Y_j + (\sigma - 1)\log(\theta_j^M P_j^C) \tag{A.4}$$

The propensity to export, meanwhile, depends on the productivity of the origin market and its factor payment w_i.

$$\begin{aligned}\delta_i^{X,M} &= (\sigma^M - 1)\log \bar{\eta}_i^M + (1 - \sigma^M)\log w_i \\ &= (\sigma^M - 1)\log \bar{\eta}_i^M + (1 - \sigma^M)\log(w_i M_i) - (1 - \sigma^M)\log(M_i)\end{aligned}$$

The factor payment w_i and labor supply M_i are co-determined. I here describe the algebra involved in eliminating the endogenous wage from the equation. In the style of Redding and Venables (2004), I define an industry-specific market access term which summarizes the relative openness of economy i outward.

$$\begin{aligned}MA_i^{X,M} &= \log \sum_{j=1}^{J} \left(\frac{T_{ij}}{\theta_{ij}^M}\right)^{(1-\sigma^M)} (P_j^M)^{\sigma^M - \sigma}(\theta_j^M P_j^C)^{\sigma-1} Y_j \\ &= \log \sum_{j=1}^{J} \exp\left[\delta_j^{M,M} + (1 - \sigma^M)\log \frac{T_{ij}}{\theta_{ij}^M}\right]\end{aligned} \tag{A.5}$$

The term under summation in the first line includes the denominator to the share λ_j^h. In the second line, I substitute the definition of $\delta_j^{M,M}$ the importer fixed effect in the first-stage regression, Equation (15). This substitution means that one can directly calculate the exporter market access from the results of the first-stage regression. Reorganizing the terms in $\delta_i^{X,M}$,

$$
\begin{aligned}
\delta_i^{X,M} &= (\sigma^M - 1)\log\bar{\eta}_i + (1-\sigma^M)\log\big[\sum_{j=1}^{N} E_j^M\big] - (1-\sigma^M)\log(M_i) \\
&= (\sigma^M - 1)\log\bar{\eta}_i + (1-\sigma^M)\log\big[\sum_{j=1}^{N} \lambda_{ij}^M Y_j\big] - (1-\sigma^M)\log(M_i) \\
&= (\sigma^M - 1)\log\bar{\eta}_i + (1-\sigma^M)\log\big[\sum_{j=1}^{N} Y_j (P_j^M)^{(\sigma^M-\sigma)}(P_j^C)^{(\sigma-1)}\left(\frac{w_i T_{ij}}{\bar{\eta}_i\theta_{ij}^M}\right)^{(1-\sigma^M)}\big] - (1-\sigma^M)\log(M_i) \\
&= 2(\sigma^M - 1)\log\bar{\eta}_i + (\sigma^M-1)\log(M_i) - (\sigma^M-1)MA_i^{X,M} - (\sigma^M-1)\log(w_i) \\
&= 2(\sigma^M - 1)\log\bar{\eta}_i^M + (\sigma^M-1)\log(M_i) - (\sigma^M-1)MA_i^{X,M} - (\sigma^M-1)\log\frac{w_i M_i}{M_i}
\end{aligned}
$$

In the first line, I apply the market clearing condition to eliminate the endogenous wage. In the last line, note that the wage has reappeared. This wage can be eliminated by following the same chain of algebra. Then, the exporter fixed effect is the sum of an infinite geometric series with radius $(1-\sigma^M)$, converging to:

$$
\delta_i^{X,M} = \frac{\sigma^M - 1}{\sigma^M}\log\bar{\eta}_i^M + \frac{(\sigma^M-1)}{\sigma^M}\log(M_i) - \frac{(\sigma^M-1)}{\sigma^M}MA_i^{X,M} \tag{A.6}
$$

Therefore, the model predicts larger exporters have larger manufacturing labor forces (M_i) and (counterintuitively perhaps, but aligned with previous work such as Duranton, Morrow, and Turner (2014)) have lower exporter market access.

A similar procedure yields the following identities for the agricultural industry.

$$
\begin{aligned}
\log E_{ij}^A &= (1-\sigma^A)\log\frac{T_{ij}}{\theta_{ij}^A} + \delta_i^{X,A} + \delta_j^{M,A} \\
MA_i^{X,A} &= \log\sum_{j=1}^{J}\left(\frac{T_{ij}}{\theta_{ij}^A}\right)^{(1-\sigma^A)}(\theta_j^A P_j^A)^{\sigma^A-\sigma}(P_j^C)^{\sigma-1}Y_j \\
&= \log\sum_{j=1}^{J}\exp\left[\delta_j^{M,A} + (1-\sigma^A)\log\frac{T_{ij}}{\theta_{ij}^A}\right] \\
\delta_i^{X,A} &= \frac{(\sigma^A-1)}{\sigma^A}\log(Q_i^A) - \frac{(\sigma^A-1)}{\sigma^A}MA_i^{X,A}
\end{aligned}
$$

In this last equation, note that the agricultural production function Q_i^A has a nonlinear dependence on unobservable land productivity $\xi(\omega)$ from Equation (11)

$$Q_i^A = \sum_{\omega=1}^{N_i} \eta^A(\omega) \frac{\exp(\gamma\pi^A(\omega) + \gamma\xi(\omega))}{1 + \exp(\gamma\pi^A(\omega) + \gamma\xi(\omega))}$$

Thus, unlike in the manufacturing regression where the error term is linear, here, errors and observables enter in a nonlinear fashion. Thus, I require a rich set of nonlinear controls and flexible functions of my choice of instrument.

On the manufacturing propensity to import in (A.4), we cannot derive a straightforward second-stage regression which identifies σ^M. The propensity to import depends jointly on the inner-nest between-country decision and outer-nest between-sector decision (as evidenced by the price index P_j^C).

$$MA_i^{M,M} = \log \sum_{j=1}^{N} \exp\left[\delta_i^{X,M} + (1-\sigma^M) \log \frac{T_{ij}}{\theta_{ij}^M}\right]$$

Then, using (14), it follows that:

$$MA_i^{M,M} = (1-\sigma^M) \log P_j^M$$

E.2 First stage estimation results

Table A.15: First stage regression: agricultural sector and quantities of trade.

	(1)	(2)	(3)	(4)
log(Population-weighted distance)	-2.065***	-2.597***	-2.580***	-2.805***
	(0.0270)	(0.0592)	(0.0601)	(0.0931)
log(Population-weighted distance)^2		0.2134***	0.1567***	0.1542***
		(0.0218)	(0.0519)	(0.0511)
log(Population-weighted distance)^3			0.0184	0.1426***
			(0.0154)	(0.0422)
log(Population-weighted distance)^4				-0.0362***
				(0.0114)
R^2	0.64420	0.64547	0.64549	0.64562
Observations	28,100	28,100	28,100	28,100
Mean effect	-0.2307	-0.2386	-0.2405	-0.2457
Median effect	-0.2599	-0.2750	-0.2784	-0.2815
Mean $(1-\sigma)\log(T_{ij})$	-3.500	-3.558	-3.572	-3.605
SD $(1-\sigma)\log(T_{ij})$	1.467	1.466	1.467	1.475
Exporter fixed effects	✓	✓	✓	✓
Importer fixed effects	✓	✓	✓	✓

NOTES: Implements first-stage regression in (15) using the CEPII BACI trade flow database from 2002-2019. Agricultural codes are HS codes 01–15 as in Tombe (2015). Population-weighted distance comes from the CEPII gravity database. Mean effects report the implied distance elasticity of trade quantities. The average value of $(1-\sigma)\log T_{ij}$ is given by the average of fitted values.

Table A.16: First stage regression: manufacturing sector and quantities of trade.

	(1)	(2)	(3)	(4)
log(Population-weighted distance)	-2.011***	-2.426***	-2.302***	-2.051***
	(0.0209)	(0.0520)	(0.0554)	(0.0771)
log(Population-weighted distance)^2		0.1591***	-0.1781***	-0.1597***
		(0.0183)	(0.0469)	(0.0463)
log(Population-weighted distance)^3			0.1045***	-0.0396
			(0.0128)	(0.0351)
log(Population-weighted distance)^4				0.0404***
				(0.0091)
R^2	0.76570	0.76631	0.76680	0.76694
Observations	36,625	36,625	36,625	36,625
Mean effect	-0.2753	-0.2842	-0.2963	-0.2891
Median effect	-0.3092	-0.3240	-0.3456	-0.3414
Mean $(1-\sigma)\log(T_{ij})$	-3.819	-3.891	-3.989	-3.934
SD $(1-\sigma)\log(T_{ij})$	1.584	1.572	1.570	1.570
Exporter fixed effects	✓	✓	✓	✓
Importer fixed effects	✓	✓	✓	✓

NOTES: Implements first-stage regression in (15) using the CEPII BACI trade flow database from 2002-2019. Manufacturing codes are HS codes 16–24 or 28–97 as in Tombe (2015). Population-weighted distance comes from the CEPII gravity database. Population-weighted distance comes from the CEPII gravity database. Mean effects report the implied distance elasticity of trade quantities. The average value of $(1-\sigma)\log T_{ij}$ is given by the average of fitted values.

Table A.17: First stage regression: agricultural sector and value of trade.

	(1)	(2)	(3)	(4)
log(Population-weighted distance)	-1.827***	-2.317***	-2.283***	-2.444***
	(0.0228)	(0.0517)	(0.0524)	(0.0805)
log(Population-weighted distance)^2		0.1968***	0.0772*	0.0754*
		(0.0187)	(0.0453)	(0.0452)
log(Population-weighted distance)^3			0.0387***	0.1279***
			(0.0133)	(0.0378)
log(Population-weighted distance)^4				-0.0260***
				(0.0101)
R^2	0.68498	0.68633	0.68643	0.68652
Observations	28,100	28,100	28,100	28,100
Mean effect	-0.1977	-0.2048	-0.2087	-0.2123
Median effect	-0.2219	-0.2353	-0.2420	-0.2442
Mean $(1-\sigma)\log(T_{ij})$	-3.255	-3.378	-3.444	-3.499
SD $(1-\sigma)\log(T_{ij})$	1.454	1.435	1.433	1.434
Exporter fixed effects	✓	✓	✓	✓
Importer fixed effects	✓	✓	✓	✓

NOTES: Implements first-stage regression in (15) using the CEPII BACI trade flow database from 2002-2019. Agricultural codes are HS codes 01–15 as in Tombe (2015). Population-weighted distance comes from the CEPII gravity database. Mean effects report the implied distance elasticity of trade values. The average value of $(1-\sigma)\log T_{ij}$ is given by the average of fitted values.

E.3 Nonhomothetic outer nest

In the outer nest, consumers trade off between agriculture and manufacturing. I introduce alternate, nonhomothetic preferences over these sectors in this section. Nonhomotheticity is required to match Engel's Law, whereby expenditure shares on food fall in income. A consumer x (where I use index x as consumers can be either landowners indexed by ω or homogeneous workers indexed by country i) with income y_x faces discrete outer nest preferences over consumption which determine the share of their income y_x devoted to sector h according to utility:

$$u^h(y_x) = (1-\sigma^h(y_x))\log P_j^h + (1-\exp(\sigma_0))\log\theta_j^h + u_{xj}^h$$

θ_j^h is a preference shifter, u_{xj}^h is an idiosyncratic preference shifter, and $\sigma_0, \sigma^h(y_x)$ are parameters I discuss next. Define

Table A.18: First stage regression: manufacturing sector and value of trade.

	(1)	(2)	(3)	(4)
log(Population-weighted distance)	-1.710***	-2.047***	-1.924***	-1.688***
	(0.0175)	(0.0464)	(0.0509)	(0.0665)
log(Population-weighted distance)^2		0.1291***	-0.2030***	-0.1856***
		(0.0159)	(0.0431)	(0.0421)
log(Population-weighted distance)^3			0.1029***	-0.0326
			(0.0114)	(0.0304)
log(Population-weighted distance)^4				0.0380***
				(0.0076)
R^2	0.80725	0.80775	0.80835	0.80850
Observations	36,625	36,625	36,625	36,625
Mean effect	-0.2118	-0.2184	-0.2294	-0.2233
Median effect	-0.2364	-0.2472	-0.2659	-0.2622
Mean $(1-\sigma)\log(T_{ij})$	-3.319	-3.377	-3.473	-3.422
SD $(1-\sigma)\log(T_{ij})$	1.378	1.365	1.359	1.361
Exporter fixed effects	✓	✓	✓	✓
Importer fixed effects	✓	✓	✓	✓

NOTES: Implements first-stage regression in (15) using the CEPII BACI trade flow database from 2002-2019. Manufacturing codes are HS codes 16–24 or 28–97 as in Tombe (2015). Population-weighted distance comes from the CEPII gravity database. Mean effects report the implied distance elasticity of trade quantities. The average value of $(1-\sigma)\log T_{ij}$ is given by the average of fitted values.

$$\log \sigma^h(y_x) = \sigma_0 + \sigma_y^h y_x + \sigma_v^h v_x,$$

where v_x is an idiosyncratic individual coefficient of substitution and σ_0 is the average substitutability of goods across consumers, akin to σ in the main text preferences. Two new parameters enter, however. First, σ_y^h represents a nonhomotheticity parameter within sector h. When $\sigma_y^h < 0$, consumers will substitute away from good h more readily as their income rises. Second, σ_v^h controls preference heterogeneity across consumers. Under this preference structure, I assume consumers agree on θ_j^h, the quality of a given sector, and that nonhomotheticity is purely driven by a price elasticity (not a quality elasticity, e.g., Handbury 2021).

I assume u_{xj}^h is distributed Type I extreme value. Then, an individual has the following expenditure share on goods in nest h:

$$\lambda^h(y_x) = \frac{(P_j^h)^{1-\sigma^h(y_x)}(\theta_j^h)^{1-\exp(\sigma_0)}}{\sum_{h\in\{M,A\}}(P_j^h)^{1-\sigma^h(y_x)}(\theta_j^h)^{1-\exp(\sigma_0)}}$$

Then, the nonhomotheticity parameter controls the price-sensitivity of expenditure shares across the income distribution.

Inclusive value. Utility across sectoral goods is given by the ex ante expected utility, or inclusive value:

$$U_j(y_x) = \frac{y_x}{\exp(\sigma_0)} \log \sum_{h\in\{M,A\}} \left(P_j^h\right)^{(1-\sigma^h(y_x))}(\theta_j^h)^{1-\exp(\sigma_0)}$$

Price index. Following Baqaee and Farhi (2019) and Nath (2022), I calculate the aggregate price index of consumption using a Törnqvist price index, which aggregates sector-level price indices P_j^h by consumer-level expenditure shares $\lambda_j^h(y_x)$,

$$P_j^C(y_x) = \Pi_{h\in\{A,M\}}(P_j^h)^{\lambda_j^h(y_x)}.$$

Aggregation across consumers. Aggregating across consumers in country j, the share of total income spent on agricultural goods is given by an average across all of the landowners and workers in j:

$$\lambda_j^h = \frac{L_j}{\Pi_j + L_j w_j}\lambda^h(w_j)w_j + \frac{\Pi_j}{\Pi_j + L_j w_j} \times \frac{1}{\Pi_j}\sum_{\omega=1}^{N_j}\lambda^h(y(\omega; s_0))y(\omega; s_0)$$

Choice of preferences. I adopt this reduced form preference structure because it is analytically convenient and nests three simpler sets of preferences in clear channels: a homothetic mixed logit ($\sigma_y^A = \sigma_y^M = 0$), a nonhomothetic but homogeneous logit ($\sigma_v^A = \sigma_v^M = 0$), and a homothetic, simple logit preference ($\sigma_v^h = \sigma_y^h = 0$). Thus, these preferences are attractive for decomposing the effects of heterogeneity and nonhomotheticity, while maintaining tractability of aggregate welfare. These preferences are closest to the preferences proposed by Fieler (2011). In principle, a related fully micro-founded preference can be found in Comin, Lashkari, and Mestieri (2021).

One could argue the inner nest should also have greater richness. Assuming a constant elasticity of substitution across sectoral goods is restrictive. However, with 114 distinct markets, the Armington assumption on goods implies 114 products in each market. Introducing greater flexibility thus runs into the curse of dimensionality without placing further structure on substitution patterns. With this added complexity, because the outer nest summarizes empirically more important channels in controlling the response to a global-scale Pigouvian tax, I maintain a simple inner nest.

Identification. Identifying $\sigma^h(y_x)$ requires data on trade shares, the price indices identified by the lower nest P_j^h, and data on incomes. I use GDP per capita from the Penn World Table to get average income in each country. I draw v_x from a standard normal distribution $v_x \sim \mathcal{N}(0,1)$. Then, for a guess of parameters $(\sigma_0, \sigma_y^h, \sigma_v^h)$, I calculate the following moment conditions:

$$m(\sigma_0, \sigma_y^h, \sigma_v^h; Z_j) = \mathbb{E}\Big[Z_j[\log(\lambda_j^A/\lambda_j^M) - (1-\sigma^A(y_x))\log(P_j^A) + (1-\sigma^M(y_x))\log(P_j^M)]\Big] \tag{A.7}$$

Using these moment conditions, σ_0 is identified by the covariance between relative shares and relative sectoral prices (the first moment condition). I use an instrument Z_j to deal with endogeneity. Shares and aggregate price indices are determined simultaneously in equilibrium, requiring the use of the instrument Z_j. For Z_j, I use a Hausman instrument of leave-one-out prices $Z_j = \frac{1}{J-1}\sum_{i\neq j, i=1}^{J} P_i^A$, which leverages variation coming from other countries' realizations of P_i^A to partial out the effects of demand shocks within country j. Identification thus relies on other countries' realizations of P_i^A being unaffected by demand shocks in a given

country $j \neq i$. σ_y^h is identified by the covariance between relative shares and income, holding fixed prices.

Estimates of the outer nest parameters are shown in A.19 with preferred estimates in Column (1). The average value of σ in agricultural industries is 0.8 and the average value of σ in manufacturing is 1.06. Relative to homothetic preferences, the results suggest that agricultural consumption falls as income rises, consistent with Engel's law (Comin, Lashkari, and Mestieri 2021; Herrendorf, Rogerson, and Valentinyi 2013). This large relative income effect implies that agricultural is less substitutable with manufacturing than a Cobb-Douglas preference might predict ($\sigma = 1$) at low incomes. Absent non-homotheticity, a heterogeneous logit (or analogous CES) slightly overpredicts substitutability across sectors with 0.92.

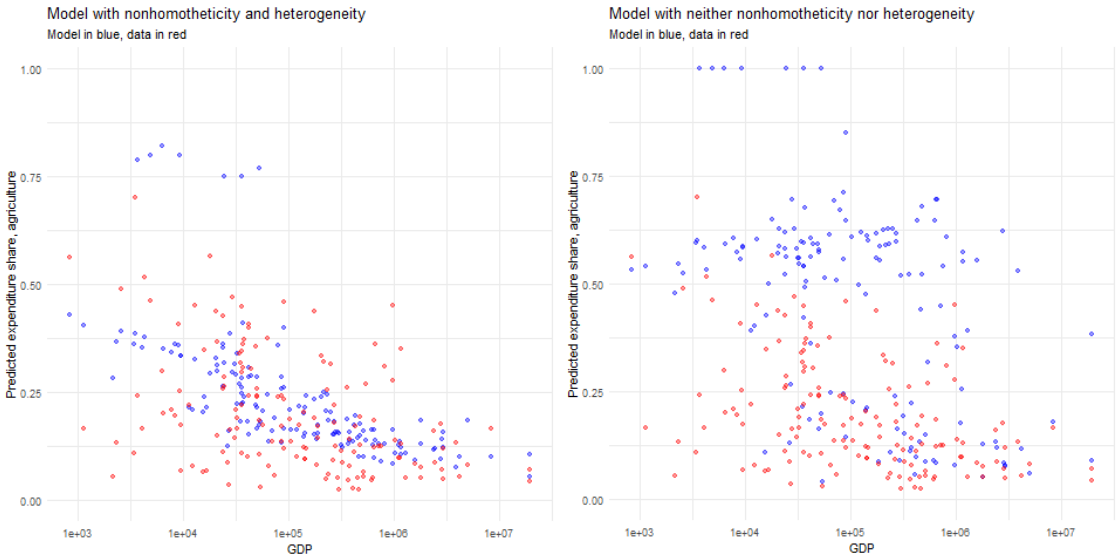
Appendix figure A.9 illustrates the correlation between GDP and the expenditure share in agriculture, λ_j^A, in the data and according the prediction of the demand-side model. I compare my data and the model using the 2017 vintage of the International Comparison Program from the World Bank (close to 2016, the final year of my deforestation test data). The fully homothetic preference, in Subfigure A.9b, completely fails to capture the negative correlation in the data. Non-homotheticity enables the model to do this as seen in Subfigure A.10c. Finally, heterogeneity may enable the data to better capture the variance around the GDP trend in Subfigure A.9a, though empirically the model rejects its importance.

Next, I conduct a validation exercise where the model is only provided information from a base period in 2011. I then evaluate the fit of the model on expenditure shares in the final year 2017. Note that I do not calibrate preference shifters to perfectly match data in 2011. Appendix Figure A.10 evaluates the fit of the mixed logit. The model has a strong overall fit, but importantly captures the covariance of the shares with GDP. I assess the J-statistic of each of the moments used in estimation. The covariance moment is largely satisfied, while the instrumental variables moment is not perfectly satisfied for around 10 countries. These exceptions include Qatar and some sub-Saharan African countries like the Gambia, Burundi, and Swaziland.

Table A.19: Estimates of the non-homothetic outer nest.

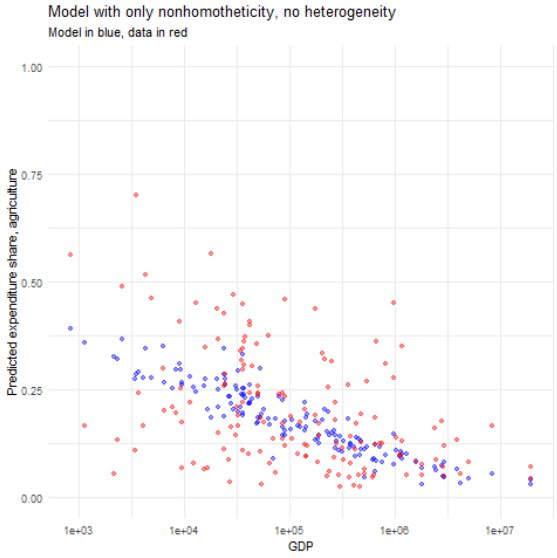
Parameter	(1) Estimate (95% CI)	(2) Estimate (95% CI)	(3) Estimate (95% CI)
		Baseline substitution parameter	
σ_0	0.06 (0.04, 0.13)	-0.08 ($-0.13, -0.02$)	0.06 ($0.05, 0.12$)
		Non-homotheticity parameter	
σ_y^A	-0.03 (-0.032, -0.27)	0 —	-0.03 ($-0.033, -0.028$)
		Consumer preference heterogeneity	
σ_v^A	0.003 (0.84, 1.11)	0.001	0 —
		Summary of $\sigma^A(y_x)$	
25th Percentile	0.78	0.92	0.78
Median	0.8	0.92	0.8
75th Percentile	0.83	0.92	0.83

NOTES: Estimates a generalized method of moments estimator for Equation (A.7). Confidence intervals are bootstrapped over 30 draws, each containing 30 countries (draws are clustered at the country level). Bootstrap confidence intervals can be asymmetric as they use the Efron (1987) bias-corrected bootstrap procedure. Data comes from the ITPD-E, where $X_j^h = \sum_{i=1}^{J} X_{ij}^h$, though estimates are consistent with tests on the World Bank International Comparison Program. Column (2) shuts down homotheticity, and column (3) shuts down preference heterogeneity.



(a) Full model.

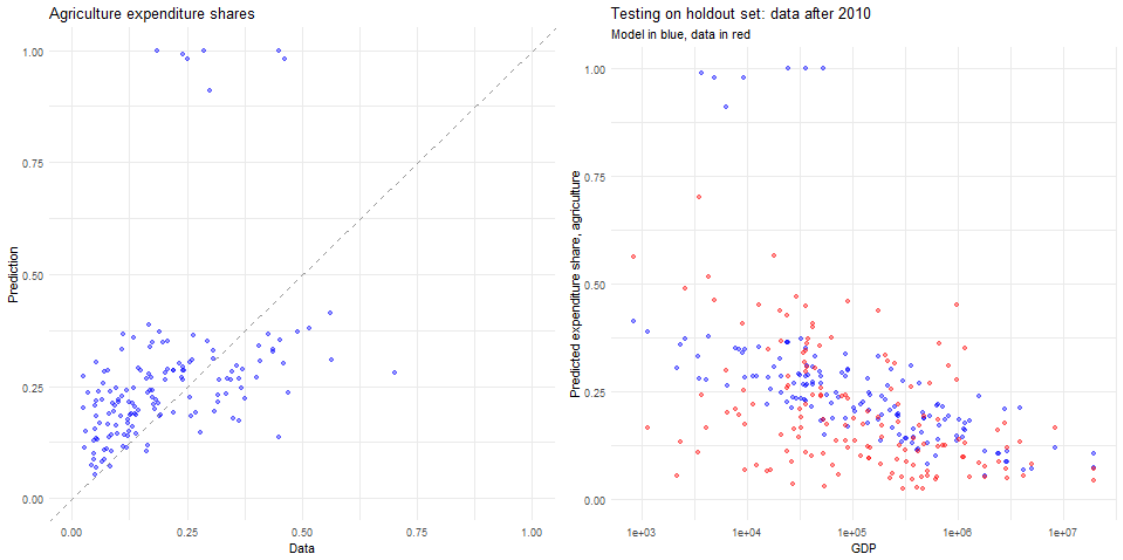
(b) Homothetic, $\sigma_v^h = \sigma_y^h = 0$.



(c) No heterogeneity $\sigma_v^h = 0$

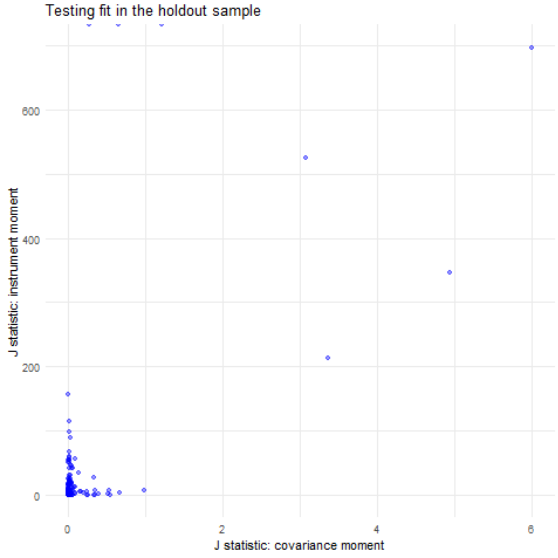
Figure A.9: Plotting the correlation between share of expenditure on agriculture and GDP (log scale).

NOTES: Agriculture expenditure data from the World Bank from the food category of the International Comparison Program. Model values come from the estimates of the GMM defined by Equation (A.7). Estimates are in column (1) of Table A.19. Observations are countries, where the data is sliced to 2019. GDP data from the Penn World Table.



(a) Fit: Overall shares.

(b) Fit: GDP moment



(c) Joint distribution of J-statistics.

Figure A.10: Evaluating model fit: training data from 1990-2009, with test set being data from 2010-2019.

NOTES: Observations in both figures are country-years. Agriculture expenditure data from the food category of the World Bank International Comparison Program. Model values come from the estimates of the GMM defined by Equation (A.7). Estimates are in column (1) of Table A.19. J-statistics computed using two-step GMM weights.

F Additional tables and figures

F.1 Metadata

Original Label	ESACCI Description	Reclassified Description	Reclassified Label
0	No data	No data	NA
10	Cropland, rainfed	Cropland	2
11	Herbaceous cover	Transition	4
12	Tree or shrub cover	Transition	4
20	Cropland, irrigated or post-flooding	Cropland	2
30	Mosaic cropland (> 50%) / natural vegetation	Cropland	2
40	Mosaic natural vegetation	Forest	1
50	Tree cover	Forest	1
60	Tree cover	Forest	1
61	Tree cover	Forest	1
62	Tree cover	Forest	1
70	Tree cover	Forest	1
71	Tree cover	Forest	1
72	Tree cover	Forest	1
80	Tree cover	Forest	1
81	Tree cover	Forest	1
82	Tree cover	Forest	1
90	Tree cover	Forest	1
100	Mosaic tree and shrub (> 50%) / herbaceous cover (< 50%)	Transition	4
110	Mosaic herbaceous cover (> 50%) / tree and shrub (< 50%)	Transition	4
120	Shrubland	Transition	4
121	Shrubland evergreen	Transition	4
122	Shrubland deciduous	Transition	4
130	Grassland	Cropland	2
140	Lichens and mosses	Transition	4
150	Sparse vegetation	Transition	4
151	Sparse tree (< 15%)	Transition	4
152	Sparse shrub (< 15%)	Transition	4
153	Sparse herbaceous cover (< 15%)	Transition	4
160	Tree cover	Forest	1
170	Tree cover	Forest	1
180	Shrub or herbaceous cover	Transition	4
190	Urban areas	Other	3
200	Bare areas	Bare	5
201	Consolidated bare areas	Bare	5
202	Unconsolidated bare areas	Bare	5
210	Water bodies	No data	NA
220	Permanent snow and ice	No data	NA

Table A.20: Reclassified labeling of the European Space Agency Climate Change Initiative data used to categorize land uses. 6 total categories were constructed, including excluded areas (NA).

Data source	Resolution	Description
European Space Agency Climate Change Initiative	300 m	Globally estimated land cover maps provided annually, 1992-2018
Food and Agriculture Organization Global Agro-ecological Zones	9 km	Actual and potential yields for a variety of crops based on climatological, physical land features (elevation, soil quality), and weather. Obtained for 2000 and 2010. In 10 kg dry weight per hectare.
Travel time to cities (Nelson et al. 2019)	1 km	a suite of nine global travel-time accessibility indicators for the year 2015 for a range of settlement size classes. Validated against Google Maps data.
World Port Index	(NA)	Shapefile of the location of world ports
City locations data (Akbar et al. 2023)	(NA)	Shapefile of major cities
World Wildlife Fund (Olson et al. 2001)	1 km	Describes bio- and eco-regions of the world (used to determine "tropical" forest regions)
Distributed Active Archive Center for Biogeochemical Dynamics, NASA	300 m	Measures aboveground biomass in Megagrams Carbon per hectare in 2010.
International Union for Conservation of Nature and Natural Resources: Red List of Threatened Species	5 km	Measures the ranged species rarity of all red list species in each cell.
Global Human Settlement Layer	30m	Maps for 1975 and 1990 on presence of built structures. Derived from Landsat image collections.
Berkeley Earth Surface Temperatures	1 arcdegree	Daily temperature minimum, maximum, and average values using a mix of sensor and interpolated data
Regrowth rates (Cook-Patton et al. 2020)	1 km	Maps carbon accumulation potential in first 30 years of secondary vegetation growth (Megagrams C ha^{-1} yr^{-1})

Table A.21: Data sources for all raster (point-cloud) data used in this project. Table lists the resolution at which each dataset is available.

Table A.22: Balance table across observations with interior land shares in 2011 and those without.

	Non-interior		Interior		Differences	
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Biomass (t/ha)	36659.0	220945.0	250211.3	289466.1	213 552.3	381.6
Yields (t/ha)	608.5	908.7	1426.2	1398.0	817.7	3.0
Population density (p/sq km)	0.4	4.0	3.8	21.6	3.4	0.0
Travel time (min)	4967.1	6097.9	660.3	1342.7	−4306.8	17.7
Regrowth rates (t/ha/yr)	3.6	1.5	2.0	1.3	−1.5	0.0
Share 0 biomass	1.0	0.2	0.0	0.0	−1.0	0.0
Share 0 yields	0.7	0.5	0.4	0.5	−0.3	0.0
Share tropical	0.0	0.1	0.3	0.5	0.3	0.0

NOTES: Interior land use grid cells must have agricultural and forested shares within $(0, 1)$ in the 2001 ESACCI data. Of 5,040,000 total grid cells, 2,366,723 are entirely ocean or water and discarded. Among the remaining terrestrial grid cells, 692,713 have interior shares of target land uses.

F.2 Model moments

Table A.23: Fit of the model, calibrated on 1982-2000 data, on 2000-2016 land use data.

Moment, 2000-2016	Model	Data
Emissions in megatons of CO_2	19 254.32	17 109.32
Correlation, emissions and potential yields	0.17	0.20
Correlation, emissions and transportation costs	−0.12	−0.10
Plot-level correlation between data and model, emissions	0.66	

NOTES: Describes the results of estimating the equilibrium of the model calibrated according to the procedure described in Section 6.1. Correlations report the Pearson correlation coefficient with 1,499,694 observations.

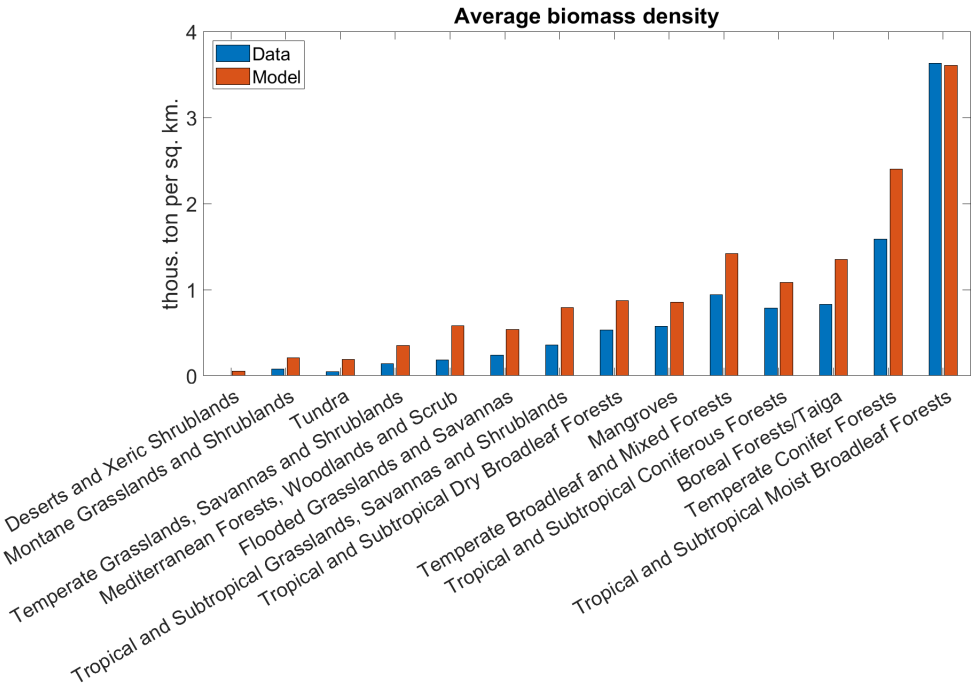


Figure A.11: Model moment: biomass density in standing forest, model vs. data, across World Wildlife Fund biomes.

NOTES: Biomass density calculated by averaging $d(\omega)$ across all grid cells in a biome.

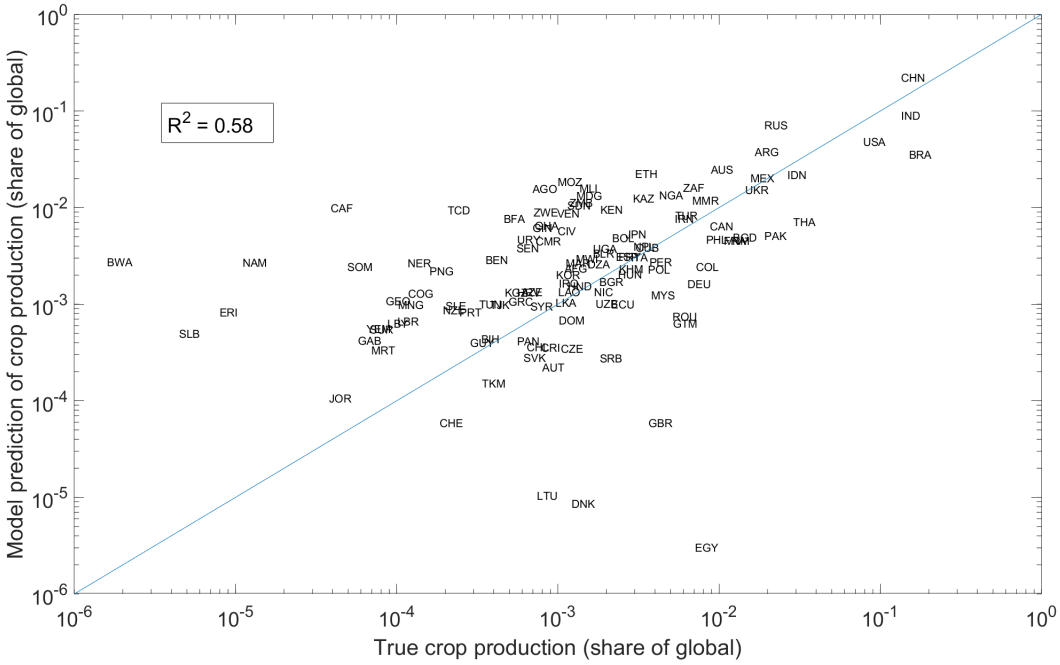


Figure A.12: Model moment: share of global agricultural production.

NOTES: Plots the share of global production which comes from each country, with the y-axis indicating the model-predicted value and the x-axis the data-implied value. Data comes from crop-country-year-level FAOSTAT production data in 2019. This production data is filtered to my crop sample, potatoes, wheat, sugarcane, rice, maize, and oilpalm. R^2 reported from a levels-on-levels regression.

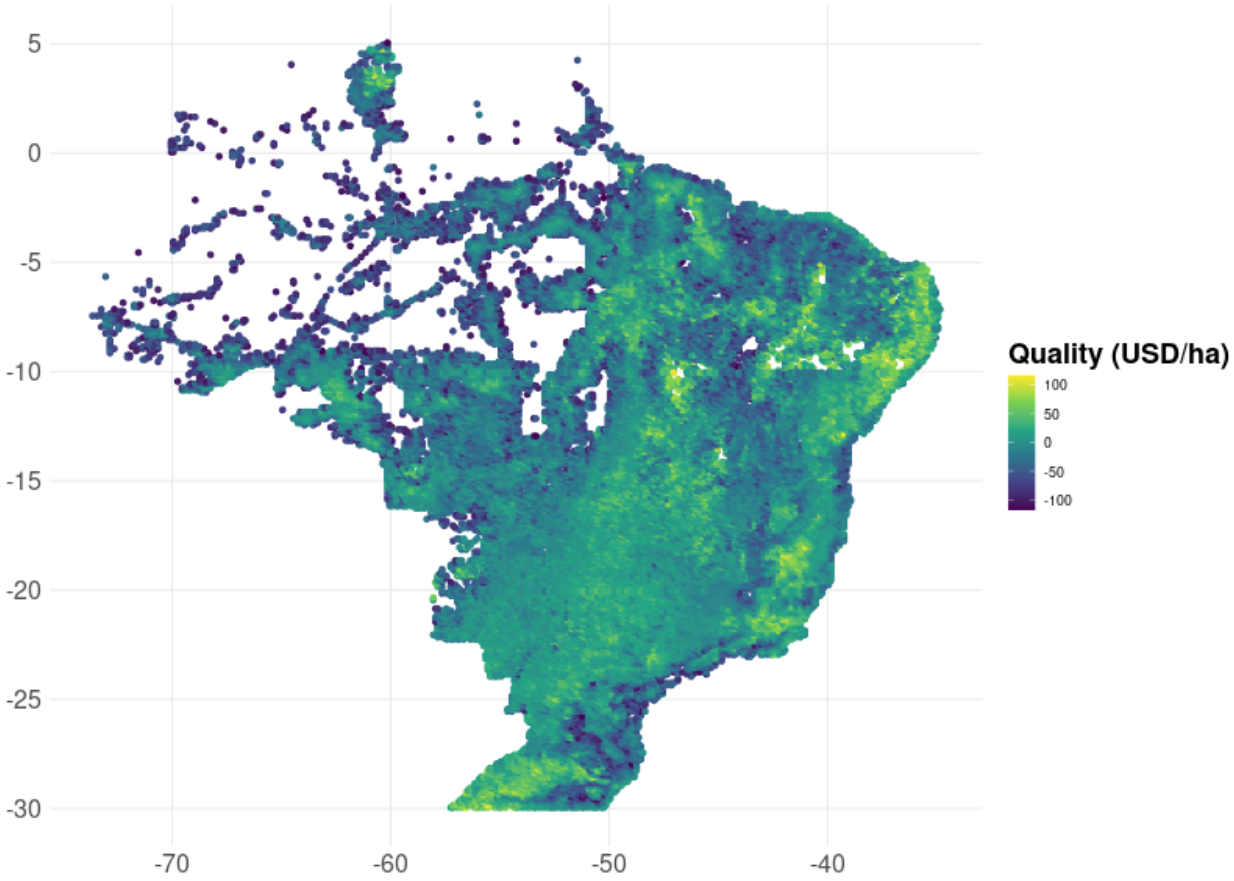


Figure A.13: Distribution of unobserved land quality shifters $\xi(\omega)$ in dollar values (divided by revenue coefficient γ) in Brazil.

NOTES: Land quality shifter calculated from the spatial first differences specification in the X-direction from column (3) of Table 2. Presents data for 2010: white/missing areas correspond to regions without interior land use shares.

F.3 Additional figures and tables: land use estimation

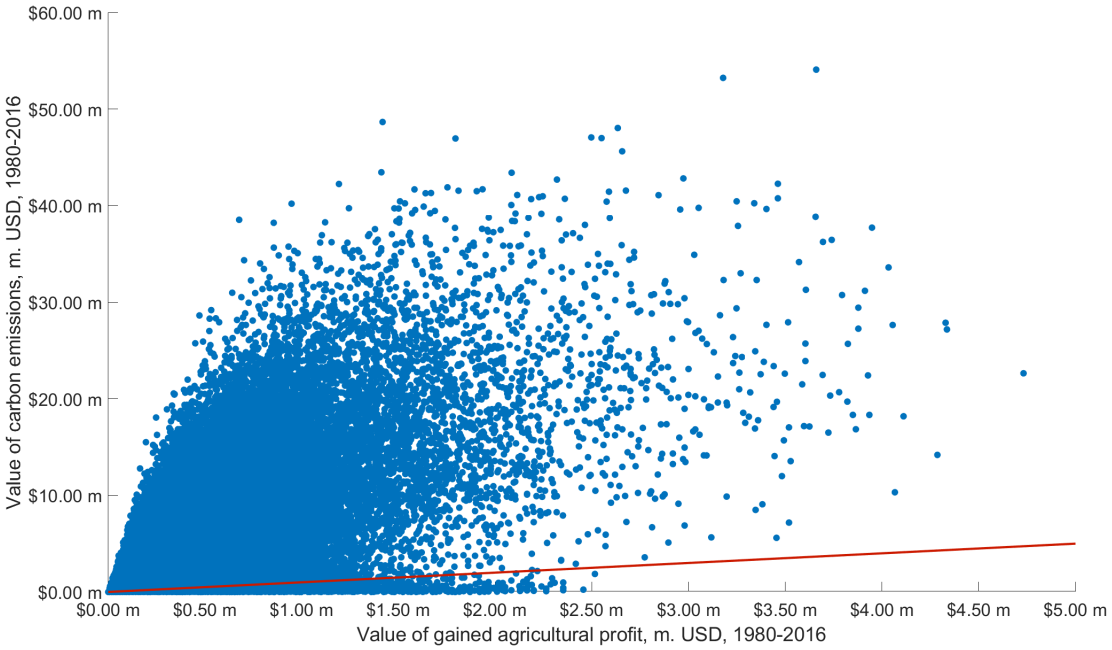


Figure A.14: Plotting environmental externality cost of deforestation against agricultural returns from deforestation, 1982-2016.

NOTES: Combines vegetation continuous fields deforestation data with the estimated γ from Table 1. Red line is the 45 degree line.

Table A.24: Linear heterogeneity extensions to baseline landowner's problem.

	LHS = Relative land use share of agriculture				
Coefficient	(1)	(2)	(3)	(4)	(5)
γ, revenue	1.606***	1.514***	1.539***	1.349***	1.925***
	(0.0352)	(0.0386)	(0.0360)	(0.0379)	(0.0365)
$c(\omega)$, travel time	-0.0240***		-0.0209***	-0.0153***	-0.0150***
	(0.0005)		(0.0005)	(0.0006)	(0.0006)
ϕ^{FA}, lagged forest share		-7.286***	-7.283***	-7.309***	-8.057***
		(0.0185)	(0.0183)	(0.0190)	(0.0267)
Interaction: revenue × travel time			-0.2029***	-0.1834***	-0.3054***
			(0.0106)	(0.0107)	(0.0114)
R^2	0.49124	0.48271	0.47911	0.48938	0.49322
Observations	5,138,107	5,138,107	5,138,107	5,138,107	5,138,107
country fixed effects	✓	✓	✓	✓	✓
accessing partner fixed effects	✓	✓	✓	✓	✓
ϕ^{FA} × country	✓				
$c(\omega)$ × country		✓			
regrowth rate × country				✓	
biomass (scrap) × country					✓

NOTES: Regressions implement 11. Standard errors clustered at the 50 km block level. Columns interact a country fixed effect with a specified cost. Interaction and revenue coefficients are both estimated with respect to the WTO accession instrument.

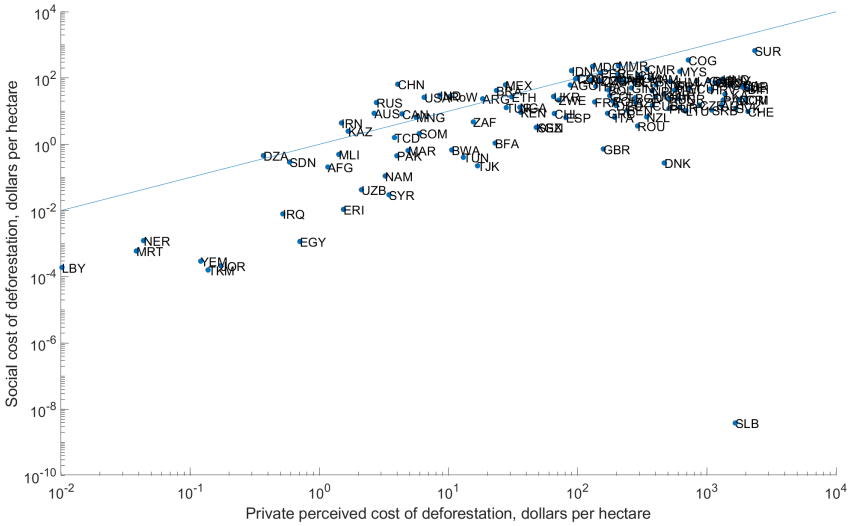


Figure A.15: Comparing the implied average switching cost in each country to its socially efficient price on deforestation.

NOTES: Calculated at the value of ϕ^{FA} from Table 1. Social cost of carbon held at \$190/ton. Uses land use data (forest cover and biomass) from 1980.

Figure A.16: Cross-country tax incidence on landowners.

NOTES: Average landowner returns calculated based on Equation (9), using FAO data on value-added per hectare to convert real baseline prices $P_i^C(0)$ to dollar terms. Omits Egypt (which has poor FAO coverage), Suriname (which is almost entirely forested), Switzerland, and New Zealand (the latter two are pure outliers).

F.4 Additional figures: Pigouvian tax counterfactual

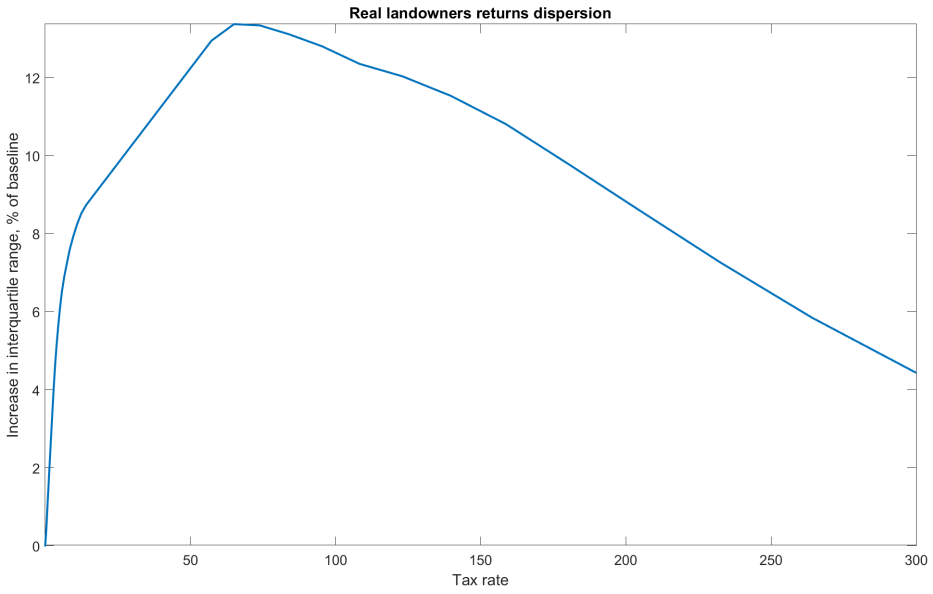
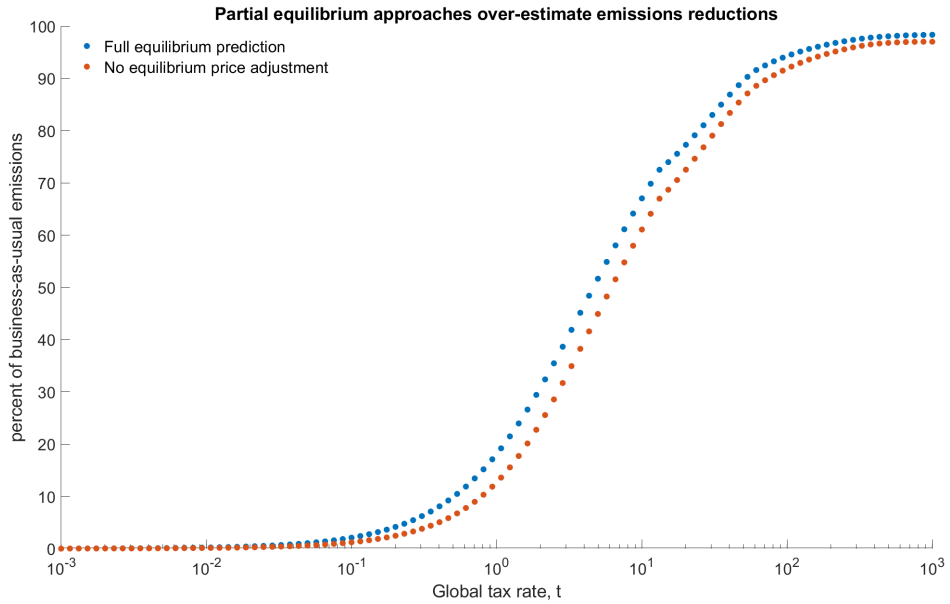
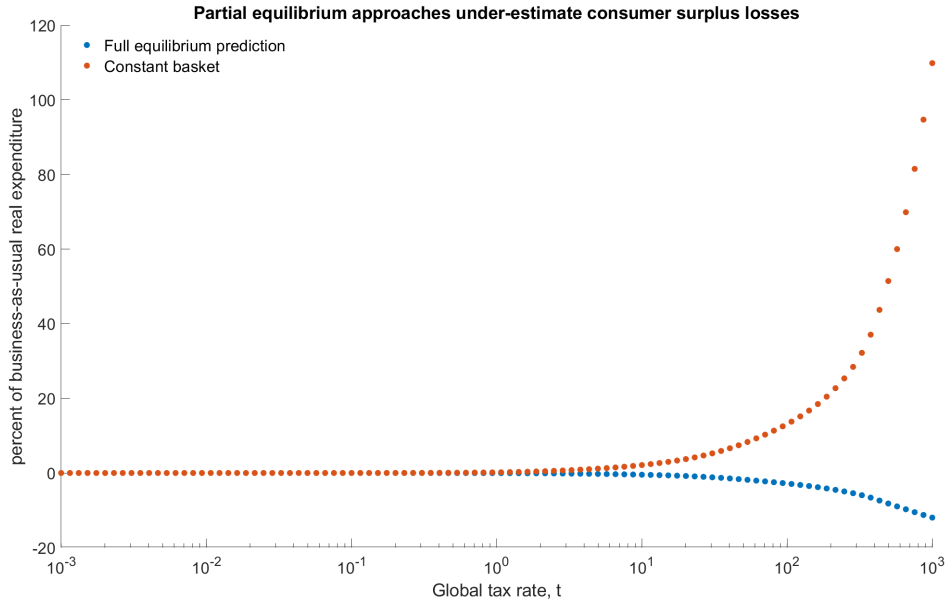


Figure A.17: Plotting the percent change in the average within-country interquartile range of landowners' profits due to global carbon pricing.

NOTES: Simulates the Pigouvian tax counterfactual for a logarithmically spaced grid of tax rates (social costs of carbon) from $[10^{-3}, 10^{3}]$. I zoom into the region from $[0, 300]$ to highlight nonlinearity in this region. Simulations assume parameters (γ, σ) are at the point estimates from Table 1. Calculates percent change in the IQR of returns within each country and aggregates using an area-weighted average.



(a) In general equilibrium, rising prices raise returns to agriculture.



(b) Consumer surplus losses in general equilibrium are larger as consumers substitute away from their preferred, but high-emissions, baskets.

Figure A.18: Illustrating the two channels which raise global equilibrium costs of abatement relative to partial equilibrium estimates.

NOTES: Simulates the Pigouvian tax counterfactual for a logarithmically spaced grid of tax rates (social costs of carbon) from $[10^{-3}, 10^{3}]$. Simulations assume parameters (γ, σ) are at the point estimates from Table 1. Subfigure A.18a shows the change in emissions avoided by a Pigouvian tax when prices and wages are held at their business-as-usual values. Subfigure A.18b shows the change in real expenditures if aggregate prices are held at their business-as-usual values.

Table A.25: Results for global Pigouvian tax of $190 using calorie-weighted sums of yields, rather than the first principal component.

Input assumption	High-input	Low-input
Emissions	0.02	0.03
(prop. of BAU)	(0.02, 0.02)	(0.03, 0.04)
Agriculture	0.99	0.91
(prop. of BAU)	(0.98, 1.00)	(0.89, 0.92)
Manufacturing	1.01	1.02
(prop. of BAU)	(1.01, 1.01)	(1.02, 1.02)
Cropland area	0.92	0.85
(prop. of BAU)	(0.92, 0.92)	(0.85, 0.86)
Welfare	0.97	0.94
(prop. of BAU)	(0.97, 0.97)	(0.94, 0.95)
Prices	1.15	4.23
(PTR, %)	(1.09, 1.20)	(4.00, 4.74)
Wages	0.11	0.30
(PTR, %)	(0.10, 0.12)	(0.26, 0.38)

NOTES: Input assumption refers to the "input assumption" column of the FAO GAEZ potential yield data. High-input refers to an assumption that crops receive fertilizer, pesticide, and drought-adaptive inputs, whereas low-input farms tend to be more susceptible to weather variation. Prop. of BAU indicates the change in quantity, in proportional terms, to the business-as-usual equilibrium ($\hat{Q} = Q^{TAX}/Q^{BAU}$). PTR % indicates pass-through as a percent of the $190 tax rate. Standard errors indicate bias-corrected bootstrap draws (Efron 1987) with 100 draws. Each draw pulls from the sampling distribution of $(\vec{\gamma}, \sigma^M)$ reported in Table 1 with $\sigma = 0.5$ and calibrated values of θ_i to match World Bank ICP consumption data.

Table A.26: Results for global Pigouvian tax of $190 with calibrated endogenous quality shifters ξ_i.

	Country-level quality shifters, ξ_i
Emissions	0.08
(prop. of BAU)	(0.06, 0.11)
Agriculture	0.94
(prop. of BAU)	(0.91, 0.97)
Manufacturing	2.33
(prop. of BAU)	(2.21, 2.64)
Cropland area	0.69
(prop. of BAU)	(0.69, 0.71)
Welfare	0.92
(prop. of BAU)	(0.92, 0.93)
Prices	21.71
(PTR, %)	(16.35, 34.13)
Wages	0.72
(PTR, %)	(0.51, 1.58)

NOTES: Prop. of BAU indicates the change in quantity, in proportional terms, to the business-as-usual equilibrium ($\hat{Q} = Q^{TAX}/Q^{BAU}$). PTR % indicates pass-through as a percent of the $190 tax rate. Standard errors indicate bias-corrected bootstrap draws (Efron 1987) with 100 draws. Each draw pulls from the sampling distribution of $(\vec{\gamma}, \sigma^M)$ reported in Table 1 with $\sigma = 0.5$ and calibrated values of θ_i to match World Bank ICP consumption data.

Table A.27: Results of a $190 Pigouvian tax with non-homothetic preference parameters from Table A.19.

	Estimate & 95% CI
Emissions	0.03
(prop. of BAU)	(0.02, 0.19)
Agriculture	0.97
(prop. of BAU)	(0.79, 1.69)
Manufacturing	1.00
(prop. of BAU)	(0.96, 1.01)
Prices	2.66
(PTR, %)	(1.00, 28.48)
Wages	0.46
(PTR, %)	(0.11, 9.96)

NOTES: Prop. of BAU indicates the change in quantity, in proportional terms, to the business-as-usual equilibrium ($\hat{Q} = Q^{TAX}/Q^{BAU}$). PTR % indicates pass-through as a percent of the $190 tax rate. Standard errors indicate bias-corrected bootstrap draws (Efron 1987) with 100 draws. Each draw pulls from the sampling distribution of $(\vec{\gamma}, \sigma^M, \sigma^A)$ reported in Table 1 with calibrated values of θ_i to match World Bank ICP consumption data. Uses a draw from the sampling distribution of reduced-form nonhomothetic preference parameters to calculate σ on each draw.

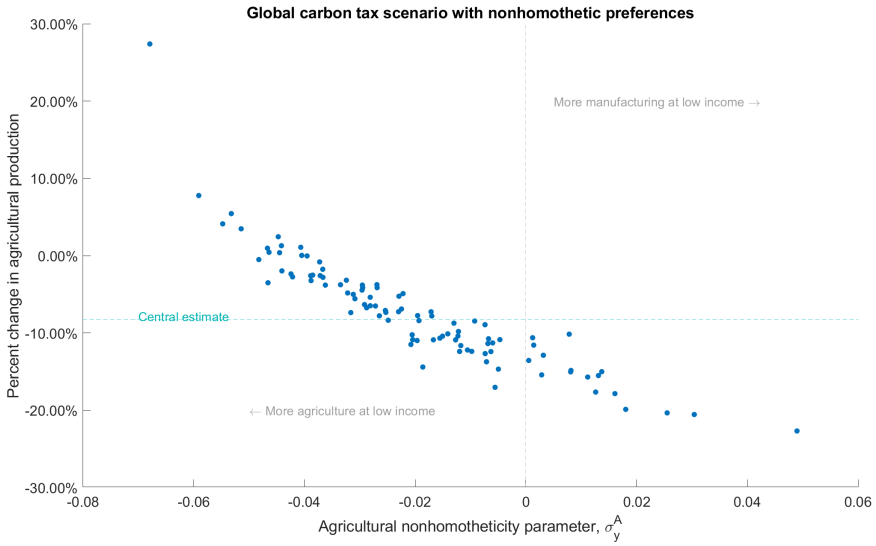


Figure A.19: Agricultural quantities in the global Pigouvian tax counterfactual with nonhomothetic preferences.

NOTES: Points each correspond to a draw of parameters from the bootstrapped confidence interval. Plots percent change in global agricultural tonnage with base being global agricultural tonnage in the business-as-usual scenario. The nonhomotheticity parameter corresponds to the reduced form nonhomothetic preference described in Appendix Section E.3. The dashed blue line highlights the central estimate at estimated values, listed in Appendix Table A.27.

F.5 Additional figures: biodiversity and co-benefits

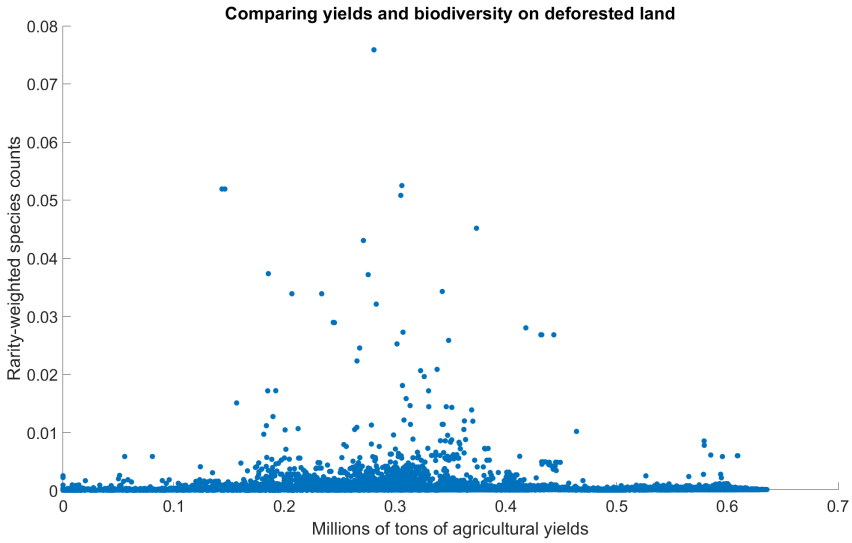


Figure A.20: Scatterplot of agricultural yields and rare species counts among land deforested since 1982.

NOTES: Points indicate one of 180,000 plots which were deforested between 1982-2016. Rarity weighted species count data comes from the IUCN Red List.

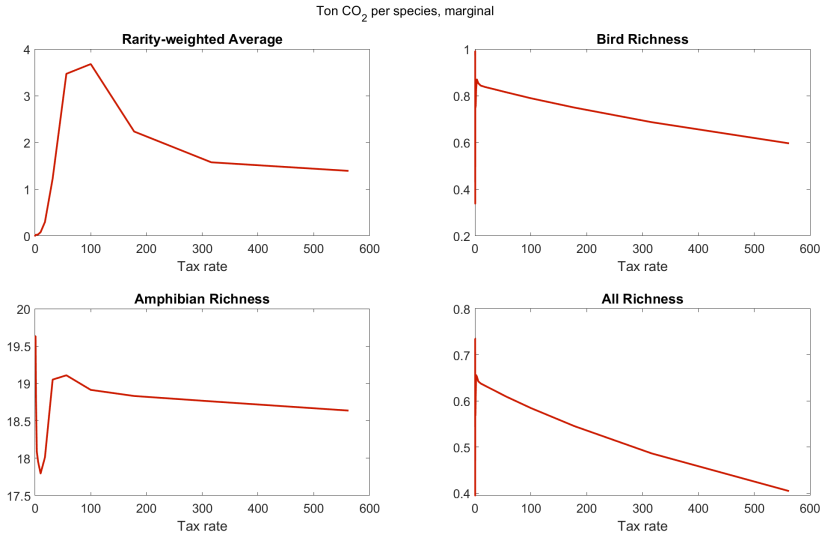


Figure A.21: Co-benefit curve: increase in emissions associated with taxing species loss.

NOTES: Interpolated curve between 25 logarithmically spaced tax rates from 0.001 to 1000. Divides changes in emissions ΔD between draws by changes in species richness measure ΔR. Assumes proportional loss of habitat from deforestation, e.g., $\Delta R = \sum_{\Omega_i} r(\omega)\hat{\mu}^A(\omega)$.