# Inappropriate Technology: Evidence from Global Agriculture*

Jacob Moscona[†] and Karthik A. Sastry[‡]

January 17, 2022

**Abstract**

An influential explanation for the persistence of global productivity differences is that frontier technologies are adapted to the conditions of the high-income, research-intensive countries that develop them and significantly less productive if used elsewhere. This paper studies how the environmental specificity of agricultural biotechnology affects its global diffusion and productivity consequences using differences in the presence of unique crop pests and pathogens (CPPs) as a shifter of the potential appropriateness of crop-specific biotechnology developed in one country and applied in another. We find that inappropriateness predicted by CPP mismatch reduces cross-country transfer of novel plant varieties and that the predicted inappropriateness of frontier technology reduces crop-specific output. Our estimates imply that this ecological mismatch reduces global agricultural productivity by 40-50% and increases productivity disparities by 10-15%. We use our framework to investigate why the Green Revolution had heterogeneous effects across environments, why adoption of frontier technology remains low in Africa, and how emergence of new R&D markets and ecological changes from global warming might affect global productivity.

*We thank Daron Acemoglu, Marios Angeletos, David Atkin, Victor Couture, Dave Donaldson, Ricardo Caballero, Arnaud Costinot, Michela Giorcelli, Michael Goodin, Gabriel Kreindler, Matt Lowe, Martí Mestieri, Paul Novosad, Nathan Nunn, Ben Olken, Jesse Perla, James A. Robinson, and Frank Schilbach for advice and comments. We are very grateful to Christine Sherratt and Hope Jansen for their help acquiring the CABI Crop Pest Compendium and to Tomochika Motomura for his support with the UPOV data. We also thank seminar participants at MIT, Harvard, UBC, the NEUDC, the RIDGE Virtual Forum, and the NBER Development Economics/BREAD Fall 2021 Meeting for helpful feedback. Miriam Zuo provided excellent research assistance.

[†]Harvard University and J-PAL, email: moscona@fas.harvard.edu; website: https://scholar.harvard.edu/moscona

[‡]MIT, email: ksastry@mit.edu; website: http://economics.mit.edu/grad/ksastry

# 1. Introduction

Research and development (R&D), which drives technological progress, is concentrated in a small set of high-income countries. The United States alone accounts for 25% of global R&D investment, and the European Union for a further 20%. By contrast, Africa and and South Asia combined account for merely 3.6%, despite encompassing 42% of the world's population (Boroush, 2020). To what extent do these vast disparities in research intensity underlie global disparities in productivity?

One school of thought starts from the premise that the most transformative technological knowledge is internationally transmittable and broadly applicable, and concludes that technology diffusion from the innovative frontier reduces global disparities and can even induce productivity convergence in the long run.[1] A second, contrasting school of thought emphasizes that much technological advancement is attuned to specific methods or factors of production (Atkinson and Stiglitz, 1969). Variations of this *inappropriate technology hypothesis* state that frontier innovators' focus on developing technology that matches local characteristics severely inhibits that technology's usefulness in, and diffusion to, other contexts (Stewart, 1978; Basu and Weil, 1998; Acemoglu and Zilibotti, 2001). In this framework, technological progress in the frontier causes productivity to persistently differ across places and cluster in those "similar" to research leaders. The quantitative relevance and global incidence of these predictions, however, remain largely unknown.

This paper empirically investigates the inappropriate technology hypothesis in a context in which all of its underlying forces loom especially large: global agriculture and plant biotechnology. Agriculture features immense and persistent cross-country productivity differences (Caselli, 2005), and global R&D is dominated by a small set of biotechnology firms in rich countries (Fuglie, 2016).[2] Despite historical recognition that this inequity may underlie productivity differences, most notably expressed in the Green Revolution of the mid 20th century, the contemporary research gap is not filled by public-sector research, just 3% of which takes place in low-income countries (Beintema et al., 2012), or philanthropically supported research, which also concentrates in wealthy countries.[3]

The core of our strategy for testing and quantifying the inappropriate technology hypothesis is a new measure of potential biotechnological inappropriateness based on the global distribution and crop-specificity of crop pests and pathogens (henceforth, CPPs). CPPs are extensively documented as pre-eminent threats to agricultural productivity and targets for biotechnological innovation (Savary et al., 2019). Our analysis exploits the fact that a given crop-country's CPP environment is a predetermined shifter of the potential effectiveness of a foreign technology originally developed for a different CPP environment. We then investigate each pillar of the inappropriate technology hypothe-

[1] Eaton and Kortum (1996) and Barro and Sala-i Martin (1997) model how free diffusion of ideas can sustain international, in growth rates and/or levels, in Neoclassical endogenous growth models. Parente and Prescott (1994, 2002) suggest that barriers to technology adoption explain an observed lack of income and growth convergence.

[2] Over 50% of private R&D occurs in North America (Fuglie, 2016), and a majority of countries in sub-Saharan Africa lack a single private sector breeding or research program (Access to Seeds Foundation, 2019).

[3] Vidal (2014), in an analysis of all grants from the Gates Foundation, find that 4% of funding for non-governmental organizations is invested in Africa, while 75% is invested in US-based organizations.

sis by studying the relationship between this determinant of appropriateness and global innovation, technology diffusion, and production. We use these estimates, interpreted via a model, to quantify the impact of disparities in research intensity and ecological mismatch on the global distribution of agricultural productivity and to study the effects of counterfactual changes to global research and ecology. In doing so, our study provides new evidence that the environment shapes comparative development. However, "better" or "worse" geographic conditions are not fixed; instead, they are determined as evolving equilibrium outcomes of endogenous technology development and diffusion.

Toward these goals, we first present a model of production and endogenous innovation in the global agricultural system. Farmers freely choose which crops to grow and what international technologies to use. Profit-maximizing innovators in each country invest research effort into improving both context-neutral attributes of technology and context-specific adaptation to country- and crop-specific environmental features, like the pest and pathogen composition. Local economies of scale, in the form of knowledge spillovers, guide innovators toward developing technology adapted to local environmental conditions and hence endogenously "inappropriate" for dissimilar environments. In the aggregate, the global production possibilities frontier is distorted toward crop-locations with environmental conditions resembling those in the most research-productive countries. We show how the strength of these effects hinges on the extent of knowledge spillovers and the relative importance of context-specific versus context-neutral components of technology. We then write the model's equilibrium conditions describing technology diffusion and production as estimable regression equations and show how to map reduced-form estimates of these equations to causal effects.

In order to directly measure the potential inappropriateness of context-specific technology across locations, we exploit the differential prevalence of crop pests and pathogens (CPPs).[4] CPPs are a dominant source of production losses, estimated to reduce annual global output by 50-80% (Oerke and Dehne, 2004). CPP resistance, and tolerance to chemicals that kill harmful CPPs, has been a key focus of traditional plant breeding (Collinge, 2016) and is central to modern transgenic crop development (Dong and Ronald, 2019). The combination of technology's CPP-specificity with large differences in CPP environments around the world can, anecdotally, limit the productivity benefit from adopting modern technology. As one example, the Maize Stalk Borer that decimates maize in Kenya is not present in the US, while the Western Corn Rootworm, nicknamed the "Billion-Dollar Bug" for its impact on US production, is not present in Kenya (Nordhaus, 2017). While the Western Corn Rootworm has been a major target for the development of resistant genetically modified varieties, the Maize Stalk Borer has received no such attention and as a result, genetically modified maize varieties are often ineffective in sub-Saharan Africa (Campagne et al., 2017).

To systematically study examples like the previous, we compile data on the global distribution and host plant species of all known CPPs—including viruses, bacteria, parasitic plants (weeds), insects, and fungi—from the Centre for Agriculture and Bioscience International's (CABI) Crop Protection

[4]Of course, the CPP environment is not the *only* characteristic that determines the direction of innovation and appropriateness of technology. In Appendix B.2 we explore the role of non-CPP differences in agro-climatic conditions.

Compendium (CPC), the "world's most comprehensive site for information on crop pests." These distribution and host plant data are based on comprehensive expert review of published literature in plant pathology, ecological science, and agronomy (Pasiecznik et al., 2005).[5] The CABI data allow us to enumerate all shared and unique CPP threats affecting any crop and country pair in the world.

We first verify the premise of the inappropriate technology hypothesis that global research is directed toward combating CPP threats present in rich countries. Using the CABI data in combination with comprehensive data on global patents that mention specific CPPs, we document that research is highly skewed toward CPPs that are present in rich, research-intensive countries. Consistent with the model's premise of home bias in CPP research, countries disproportionately patent technologies referencing locally present CPPs and this force generates the aggregate technological bias toward pathogen threats in the high-income countries where innovation takes place.

We then develop a "CPP Mismatch" measure that summarizes differences in CPP species composition at the level of crops and country pairs using techniques from population ecology (Jost et al., 2011). We use CPP Mismatch as our main measure of "potential inappropriateness" of a crop-specific technology adapted for one CPP environment and applied in another. From an empirical design perspective, this measure incorporates variation across both country pairs, which have different local CPPs, and across crops, which are host plants to different CPPs. Thus, we can conduct all subsequent analysis holding fixed differences, ecological or otherwise, purely across crops or country-pairs.

Our first main goal is to document how inappropriateness shapes global technology diffusion. We compile a unique data set on all international instances of intellectual property (IP) protection for agricultural biotechnology from the International Union for the Protection of New Varieties of Plants (UPOV), the non-governmental body tasked with codifying and administering IP protection for plant varieties. We exploit the UPOV's unique variety identifiers to track individual seed varieties from their first introduction to all other countries where they were ever transferred. We find that CPP mismatch substantially lowers cross-border transfer of technology conditional on all two-way fixed effects to absorb any average differences across country pairs or crop-specific conditions at the origin and destination. In our most conservative model, CPP dissimilarities reduce international technology transfer by 30% for the median crop and country-pair. These effects increase drastically, between six- and thirty-fold, when sub-setting to origins with more active biotechnology sectors. This result is consistent with the knowledge spillovers mechanism in the model, and it reveals the especially large technological cost of being environmentally dissimilar from frontier innovators.

Having established that CPP differences inhibit technology diffusion, our second main goal is to investigate implications for global production and specialization. Our framework predicts that countries should specialize in crops for which ecological conditions most resemble those in frontier innovating nations due to the availability of more appropriate international technology. We measure "CPP mismatch with the frontier" by either (i) imposing the United States as the single hub for global

[5]These data are commonly used in population ecology and crop science. See, for example, studies by Bebber et al. (2013), Bebber et al. (2014a), Paini et al. (2016), and Savary et al. (2019).

agricultural innovation, a fact borne out in our own technology data and consistent with others' analysis (e.g. Fuglie, 2016), or (ii) selecting the countries that develop the highest number of varieties for each crop in the UPOV certificate data. We show that countries produce less of specific crops if their local crop environment is more different from the frontier's, holding fixed country and crop effects and using a range of strategies to control directly for innate local suitability.[6] We find similar effects across regions within countries, using state-level production and CPP distribution data from India and Brazil, and using crop-level exports instead of physical output as the dependent variable. The estimated effects are large relative to observed variation in output—a one-standard deviation increase in CPP dissimilarity to the frontier reduces production of a crop by 0.51 standard deviations.

Our results so far have investigated the inappropriate technology hypothesis in a modern cross-section of all countries and crops. We next directly investigate the relationship between the appropriateness of technology and realized technology adoption, using two specific case studies in which disparities in technology adoption are the subject of intense debate. We first analyze how inappropriateness shaped the consequences of the Green Revolution of the 1960s and 1970s, perhaps the most concerted effort to shift agricultural innovative focus in history. In the Green Revolution, philanthropic organizations funded the development of breeding programs in tropical environments. We find that adoption of these new varieties and expansions of production from 1960-1980 were severely inhibited in country-crop pairs with CPP environments dissimilar to the locations of the international agricultural research centers that led research for specific crops. This supports scholars' arguments that even innovation tailored to more tropical ecosystems was not one-size-fits-all (Pingali, 2012), and directly illustrates how "advantageous" ecology changes over time as international research evolves.

Next, we study whether inappropriateness contributes to the limited use of improved agricultural inputs by smallholder farmers in Africa. Using data from the latest geo-coded round of each World Bank Integrated Survey of Agriculture (ISA), we find that farmer-crop pairs that have greater CPP mismatch with frontier countries are less likely to use improved seed varieties.[7] This suggests that features of frontier technology itself—its poor adaptation to the African environment—may explain a significant portion of farmers' low technological uptake and, by reducing potential market size, even further dissuade the development of advanced agricultural technology in the region.

Having documented each component of the inappropriate technology hypothesis, we return to our model to draw out the aggregate productivity consequences. Our calibration combines our reduced-form estimates of the effect of ecological dissimilarity on production and specialization with external estimates of the price and supply elasticities, which allow us to account for production reallocation and price effects in response to changes in the underlying productivity distribution. We

[6]These strategies include: (i) directly controlling for estimates of crop-specific potential yield in the absence of modern technology from the FAO GAEZ agronomic model, and (ii) a machine learning approach that controls flexibly for a large set of ecological features interacted with crop fixed effects, as well as CPP fixed effects accounting for the direct effect of each CPP. Our findings are consistent with historical evidence suggesting that there was nothing "special" about the innate, agro-climatic characteristics of the US and other frontier countries (Kloppenburg, 2005; Olmstead and Rhode, 2008).

[7]The ISA covers eight countries, including Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda.

first study a counterfactual scenario of "removing inappropriateness" by eliminating the knowledge gap between frontier and non-frontier CPP research. We estimate that inappropriateness reduces average global agricultural productivity by 40-50% and that losses are concentrated in Asia and Africa, underscoring the relevance of historical and current efforts to encourage biotechnological development in these neglected agricultural ecosystems. These effects explain 10-15% of cross-country disparities in productivity, driven by the fact that the countries that are most lacking in appropriate biotechnology are also those that are least productive today.

We next use our model explore how changes in the geography of innovation and ecology would affect patterns of productivity growth in three more realistic counterfactuals. In the first, we identify the countries where research investment could have the largest potential effect on global productivity after taking into account the global network of environmental mismatch. Our results convey potentially large gains from focusing a "Second Green Revolution" in India, China, and sub-Saharan Africa. In the second, we study a "BRIC realignment" that replaces the observed technological frontier with Brazil, Russia, India, and China, countries that contribute a rapidly growing share of global R&D. While far from an explicitly targeted "Second Green Revolution," this scenario is on net favorable for the world's least productive countries while harmful toward parts of Europe and North America. In the third, we study a potentially large poleward shift in the habitable range of CPPs due to climate change (Bebber et al., 2013). This ecological disruption the mismatch between countries even while holding the identity of the frontier fixed. Our results suggest that climate change could coordinate international research on a more common set of threats, and therefore that the inappropriate technology mechanism might ameliorate some of the direct productivity losses.

**Related Literature.** This paper builds on a largely theoretical body of work on the role of "appropriate technology" in shaping productivity differences (Atkinson and Stiglitz, 1969; Stewart, 1978). Early studies investigated the specificity of technical advances and barriers to their adoption within countries (Griliches, 1957; David, 1966; Salter, 1969). Stewart (1978) discusses how the inappropriateness of rich-country technology for application in low-income countries could inhibit economic development. More recent work has investigated the aggregate consequences of inappropriateness due to differences in capital intensity or skill endowment across countries (Basu and Weil, 1998; Acemoglu and Zilibotti, 2001; Caselli and Wilson, 2004; Caselli and Coleman II, 2006; Jerzmanowski, 2007). Our focus is instead on ecological differences, which perhaps cause the most acute inappropriate technology problem since the underlying differences in endowments are (essentially) immutable.

A large literature has studied the *direct* effects of adverse environmental conditions on economic development (see, for instance, Kamarck, 1976; Sachs and Warner, 1997; Gallup et al., 1999). We focus instead on how ecological mismatch affects the development and diffusion of technology. This confluence of ecology and technology diffusion is one mechanism in the theory of Diamond (1997), who discusses the easier diffusion of agricultural technology across "horizontal" landmasses. Our findings extend a more recent body of work suggesting the relationship between geography and development is endogenous, shaped by historical events and institutions (e.g., Sokoloff and Engerman, 2000; Acemoglu et al., 2001; Engerman and Sokoloff, 2002; Nunn and Puga, 2012; Alesina et al., 2013).

By proposing and quantifying a new source of productivity differences in global agriculture, we build on prior work investigating the sources of international disparities in agricultural production (e.g. Caselli, 2005; Lagakos and Waugh, 2013; Gollin et al., 2014; Adamopoulos and Restuccia, 2014). Especially related are analyses of the role of technology in shaping productivity gaps, many of which are focused on the 20th century's Green Revolution (e.g., Foster and Rosenzweig, 1996, 2004; Evenson and Gollin, 2003a,b; Pingali, 2012).

At the center our hypothesis are the determinants of technology diffusion (Keller, 2004; Kerr, 2008; Comin and Mestieri, 2014). Related work includes macro-level studies of technology diffusion in the 18th century (Comin and Hobijn, 2004, 2010; Comin and Mestieri, 2018) and micro-level studies of technology upgrading in modern times (Bandiera and Rasul, 2006; Conley and Udry, 2010; Hardy and McCasland, 2021; Verhoogen, 2021, for a review).[8] While most work in this area focuses on the characteristics of producers, our study documents how the focus of innovators determines patterns of technology adoption. Related to our hypothesis, Suri (2011) argues that differences in hybrid maize adoption in Kenya reflect differences in returns to adoption—a feature of the technology itself—and not adoption frictions.

Finally, there is a broad parallel between our analysis of ecological difference and its effects on agricultural biotechnology development and studies of globally heterogeneous human disease burdens. "Neglected Tropical Diseases" receive little attention from medical researchers in advanced economies (Kremer, 2002; Kremer and Glennerster, 2004) and inflict heavy health damages in many tropical and low-income countries (Hotez et al., 2007, 2009).

**Outline.** This paper is organized as follows. Section 2 describes a theoretical model that structures our empirical analysis and quantification. Section 3 provides background information on the ecological specificity of biotechnology and describes our measure of inappropriateness. Section 4 reports our results on international technology transfer, Section 5 reports our results on production, and Section 6 presents our findings on technology adoption. Section 7 quantifies the total effect of inappropriateness and explores counterfactual scenarios. Section 8 concludes.

# 2. Model

We first present a model of innovation, technology diffusion, and production. Relative to existing models of endogenous inappropriate technology (e.g., Acemoglu and Zilibotti, 2001), we particularly emphasize two features which are central to the context of global agriculture: the possibility for substitution across sectors and production technologies (e.g., crops and crop varieties) and the multi-dimensional nature of environmental differences. We use the model to introduce the key economic

[8] Also related to this paper is work investigating the *impacts* of technology diffusion; for example, Giorcelli (2019) and Giorcelli and Li (2021).

mechanisms of the inappropriate technology hypothesis and generate estimable equations for the effect of ecological differences on technology diffusion and production. We also return to the model structure in Section 7 in order to study counterfactual scenarios.

## 2.1 Set-up

### 2.1.1 Production

There is a set of countries indexed by $\ell \in \{1, \ldots, L\}$ and a set of crops indexed by $k \in \{1, \ldots, K\}$. In each country, there is a continuum of farms indexed by $i \in [\ell-1, \ell)$. Each farm can produce any of the $K$ crops with one of $L$ production technologies (e.g., crop varieties) indexed by its country of origin. Potential physical output of a farm producing crop $k$, with technology $\ell'$, in country $\ell$, on farm $i$ is denoted by the random variable $\psi_i(k, \ell')$:

$$\psi_i(k, \ell') = \omega(k, \ell) \cdot \theta(k, \ell' \to \ell) \cdot \varepsilon_i(k, \ell') \qquad \forall i \in [\ell - 1, \ell) \tag{2.1}$$

The first term, $\omega(k, \ell) \in \mathbb{R}_+$, captures average innate productivity for crop $k$ in country $\ell$. The second term, $\theta(k, \ell' \to \ell) \in \mathbb{R}_+$, captures the productivity of technology from $\ell'$ used in $\ell$. The third term $\varepsilon_i(k, \ell' \to \ell)$, is an idiosyncratic perturbation with a Fréchet distribution with mean one and shape parameter $\eta > 0$.[9] The random component captures un-modeled plot-level heterogeneity and disciplines the elasticity of average farmer choices to changes in innate or technological productivity.[10]

Farmers face an international price $p(k)$ for each crop $k$ and pay input costs equal to a fraction $\overline{\rho} < 1$ of revenue.[11] Each farmer in country $\ell$ observes prices and potential productivities, and chooses a crop-technology combination to maximize revenue. This discrete choice structure for production and specialization is similar to that used by Eaton and Kortum (2002), Costinot et al. (2016), and Sotelo (2020), and it will enable tractable analysis.

### 2.1.2 Ecological Characteristics and Ecologically-Specific Technology

We now introduce our notion of environmental differences and the adaptation of technology to these differences. Each location-by-crop pair is associated with a set $\mathcal{T}(k, \ell)$ of local ecological characteristics, which are normalized to have measure one.[12] These characteristics, importantly, may partially but not completely overlap between countries for a fixed crop. Consistent with our empirical analysis, we will think of $\mathcal{T}(k, \ell)$ describing all locally present crop pests and pathogens (CPPs).

[9]The normalization to mean one implies that the scale parameter is $(\Gamma(1 - \frac{1}{\eta}))^{-1} > 0$. This normalization is convenient for subsequent expressions; otherwise the scale factor would scale aggregate productivity.

[10]The specific Fréchet distributional assumption has two roles. First, it allows for simple analytical expressions for farm choices. Second, it determines the relationship between average and marginal products of land conditional on a specific use. Proposition 2 and the subsequent discussion highlight this latter property.

[11]We focus in this section on a world economy with fixed prices. It is straightforward to extend all analysis to a case in which prices are determined along a world demand curve for each crop. We use such an extended model to study counterfactual scenarios in Section 7.

[12]Note that any direct productivity effects of these characteristics can be modeled in innate productivity, $\omega(k, \ell)$.

A given technology, which is designed in country $\ell'$ for use in $\ell$ on crop $k$, is described by a context-neutral characteristic, $A(k, \ell') \in \mathbb{R}_+$, and a collection of context-specific characteristics, $(B(t, k, \ell' \to \ell))_{t \in \mathcal{T}(k,\ell)} \in \mathbb{R}_+^{|\mathcal{T}(k,\ell)|}$. These characteristics combine to determine the overall productivity of the technology in the following Cobb-Douglas way:

$$\theta(k, \ell' \to \ell) = \exp\left(\alpha \log A(k, \ell') + (1-\alpha) \int_{\mathcal{T}(k,\ell)} \log B(t, k, \ell' \to \ell)\, \mathrm{d}t\right) \tag{2.2}$$

where $\alpha \in (0,1)$ parameterizes the relative importance of the context-neutral characteristic. High $A$, by definition, boosts the productivity of technology in all locations $\ell$. Each characteristic $B(t)$, by contrast, affects productivity only if the characteristic (i.e. pest or pathogen) $t$ is present. Finally, the two components are complementary to one another: high general productivity increases the marginal value of resistance, and vice-versa.[13]

### 2.1.3 Endogenous Innovation

We finally specify how technology is produced. In each country $\ell'$, there is a continuum of symmetric innovators indexed by $j \in [\ell' - 1, \ell')$, who develop technology for each of the crops $k$ and destinations $\ell$. Each innovator produces a potentially different product, with $j$-specific general and ecological characteristics, that farmers cannot distinguish from one another *ex ante*. This structure of competition is not crucial for our main conclusions but will allow for a simpler characterization of each innovator's maximization problem without sacrificing the key market forces of interest.[14] We will focus on equilibria in which all innovators make symmetric choices.

Innovators choose the characteristics of technology maximize profits, which equal a fraction $\rho(\ell, \ell') \leq \overline{\rho}$ of their customers' revenue (e.g., net of trade and licensing costs), net of convex, additively separable research costs. We denote the costs of developing CPP-resistance technology as $C(z; t, k, \ell')$, for research level $z$, CPP $t$, crop $k$, and innovating country $\ell'$.[15] For tractability, we assume these costs have a power form, parameterized by $\phi > 0$, with a *knowledge spillover* from the (geometric) mean local research on the pest, $B(t, k, \ell' \to \ell') := \exp\left(\int_{\ell'-1}^{\ell'} \log B_j(t, k, \ell')\, \mathrm{d}j\right)$. We write the costs as:[16]

$$C(z; t, k, \ell') = \exp(-\tau(B(t, k, \ell' \to \ell'))) \cdot \frac{(B_0 z)^{1+\phi}}{1+\phi} \tag{2.3}$$

[13]One example of this "two-component" structure to agricultural research comes from the case study of wheat development at the International Maize and Wheat Improvement Center (CIMMYT) in the 1960s. Reynolds and Borlaug (2006) emphasize that the key challenge was to both improve yields by incorporating a specific semi-dwarfism trait ("$A$") and to increase resilience to damaging fungal wheat rusts ("$B$"), whose threat only *increased* as plants become more productive.

[14]The missing forces, relative to a model in which the innovative varieties are distinguishable imperfect substitutes, are innovators' internalizing the effects of their technology improvement on a given country's aggregate production mix and productivity. We argue that the present model, in which innovators act as if they have "small" impacts, is a more realistic description of incentives.

[15]The costs of general technology need not be specified to derive our main results.

[16]We will further require the technical condition $\phi > \eta - 1$ to ensure that the fixed-point equation determining technology quality has well-behaved, monotone comparative statics for any value of $\alpha$.

where $B_0 > 0$ is a constant and the function $\tau : \mathbb{R}_+ \to \mathbb{R}_+$, which we assume to be non-decreasing and to satisfy $\tau(0) = 0$, controls the knowledge spillover in units of "percentage cost reduction."

The knowledge spillover creates a local economy of scale. Abstractly, this could embody local sharing of ideas and scientific knowledge. More literally, it may embody the role of physical inputs with a public-good property like local test fields and local germplasm (genetic material). We will discuss examples of this phenomenon at length in Section 3.1.

Finally, note that only the knowledge spillover and the heterogeneity in the appropriation fraction $\rho(\ell, \ell')$ create "primitive" incentives to focus innovation on certain environmental characteristics or in certain locations. Otherwise, an agricultural innovator is free to direct its research toward whatever application (e.g., producing market, crop, and pest threat) is most economically profitable.

## 2.2 Main Predictions

In Appendix A, we include detailed derivations of the model's equilibrium conditions. Here, we highlight the main predictions which motivate our empirical analysis.

### 2.2.1 Technology Diffusion

Let $\delta(k, \ell', \ell)$ be the measure of $k$-CPPs that are not shared between locations $\ell$ and $\ell'$. Our first result describes how technology depends negatively on the ecological mismatch between locations, as summarized by $\delta(k, \ell', \ell)$:

**Proposition 1.** *Technology diffusion from country $\ell'$ to $\ell$ for crop $k$ can be expressed as*

$$\log \theta(k, \ell' \to \ell) = \beta(k, \ell') \cdot \delta(k, \ell', \ell) + \chi(k, \ell) + \chi(k, \ell') + \chi(\ell, \ell') \tag{2.4}$$

*where the $\chi(\cdot)$ are additive effects varying at the indicated level and*

$$\beta(k, \ell') = -\frac{(1-\alpha)\tau(B(k, \ell'))}{1 + \phi - (1-\alpha)\eta} \leq 0 \tag{2.5}$$

*where $B(k, \ell')$ is the extent of $(k, \ell')$ CPP research on CPPs present in $\ell'$.*

The proof in Appendix A.2 contains exactly expressions for each of the "fixed effects" as functions of economic primitives. In brief, $\chi(k, \ell)$ ("crop-by-destination") depends on the destination's market size and productivity; $\chi(k, \ell')$ ("crop-by-origin") depends on the scale of research in the innovating country; and $\chi(\ell, \ell')$ ("origin-destination") depends on the bilateral appropriability $\rho(\ell, \ell')$.

Ecological differences depress technology transfer, or $\beta(k, \ell') < 0$, only if both of the following two conditions hold: there is some context-specificity of technology ($\alpha < 1$) and some knowledge spillover ($\tau > 0$). Absent context-specific technology, innovation is biased toward the crops over-represented in large markets, but not the large-market ecological conditions for growing those crops. Absent the knowledge spillover, innovation would concentrate on large-market ecological conditions, but this would have no external effects on the rest of the world.[17] With both ingredients ($\alpha < 1$ and $\tau > 0$), by contrast, innovators in country $\ell'$ have a "knowledge gap" about local ecological characteristics relative to others and therefore produce more technology for ecologically similar destinations. A lower elasticity of supply ($\phi$) and higher elasticity of demand ($\eta$) amplify this effect.

If local knowledge spillovers scale with local research, or $\tau(B)$ is strictly increasing, then $|\beta(k,\ell')|$ increases in the sending country's CPP research intensity $B(k,\ell')$. Under this case of the model, geographic differences relative to the most active innovating countries are most costly for technology transfer and productivity. If instead knowledge spillovers were purely on the extensive margin, or $\tau(B) \equiv \tau$ for all $B > 0$, we would observe an equal marginal effect of environmental differences on technology transfer from "high-tech" and "low-tech" sending countries.

In our empirical analysis, we will estimate Equation 2.4 treating counts of uniquely identified seed varieties transferred across borders as a proxy for $\theta(k,\ell' \to \ell)$ and using our measurement of CPP mismatch as a proxy for $\delta(k,\ell',\ell)$.[18] We will also directly investigate whether the effect of environmental differences on technology transfer is exaggerated when the origin country is on the "research frontier," measured via various empirical proxies.

### 2.2.2 Specialization and Productivity

We next translate the consequences of inappropriate technology for production. A key issue that our model handles precisely is selection along unobserved dimensions of land quality. While secularly boosting the productivity of a given crop (e.g., by improving available foreign technology) moves out the production possibilities frontier in any location, it also encourages more production of that crop on relatively less-suitable land. Due to this selection effect, in a model with unobserved plot-level heterogeneity, the appropriateness of technology has ambiguous effects on measured average productivity. We will exploit our parametric assumption of Fréchet-distributed plot-level shocks to derive exact and economically interpretable predictions for observed production, planted areas, and yields, which will allow us to infer the productivity consequences of inappropriate technology.

Toward this end, we first define the *crop technology index* $\Theta(k,\ell)$ and *revenue productivity index* $\Xi(\ell)$ as a function of local technology and productivity shifters:

$$\Theta(k,\ell) = \left(\sum_{\ell'=1}^{L} \theta(k,\ell' \to \ell)^{\eta}\right)^{\frac{1}{\eta}} \qquad\qquad \Xi(\ell) = \left(\sum_{\ell'=1}^{L} \Theta(k,\ell)^{\eta}\omega(k,\ell)^{\eta}p(k)^{\eta}\right)^{\frac{1}{\eta}} \tag{2.6}$$

The following result summarizes the model predictions:

[17] In Acemoglu and Zilibotti (2001), there are no knowledge spillovers but instead "copycat producers" who replicate technologies in other countries and compete away all potential profits to the original innovator. This creates a similar uninternalized effect of home-country research on foreign production while implying, sharply, that the original inventor produces nothing in other countries and responds not at all to market-size incentives in those countries. These latter predictions are counterfactual in the context of plant biotechnology, which as we will document features extensive international research and technology transfer.

[18] We describe the measurement of each of these variables in Sections 4.1 and 3.4, respectively.

**Proposition 2.** *Production of crop $k$ in country $\ell$, $Y(k, \ell) > 0$, is given by*

$$\log Y(k, \ell) = \eta \log \Theta(k, \ell) + \eta \log \omega(k, \ell) + (\eta - 1) \log p(k) + (1 - \eta) \log \Xi(\ell) \tag{2.7}$$

Production is monotone increasing in the index of technology from each source country, and hence positive shifters of this index. In Equation 2.7, crop and country fixed effects respectively absorb (international) prices and average local revenue productivity. In the proof of this result in Appendix A.3, we derive also the model's predictions for physical yield and planted area. Because of the Fréchet model's implication that selection effects directly net out direct productivity effects, log crop-specific yields are predicted to have no relationship with measured technological inappropriateness conditional on country fixed effects.

In our empirical analysis of Section 5, we will estimate Equation 2.7 using CPP mismatch with an empirically identified technological frontier to span $\log \Theta(k, \ell)$, crop and country fixed effects to span prices and aggregate revenue productivity, and a variety of empirical strategies to span innate productivity $\omega(k, \ell)$. This will allow us to directly measure the effect of inappropriateness on production choice and specialization. We will also directly test the model's predictions for area and yields to assess the validity of the specific Fréchet model for unobserved heterogeneity.

In Section 7, we will use the estimates from this analysis plus the model structure to estimate causal effects on revenue productivity. In short, this process amounts to a "two-step strategy" of inferring the productivity effect of inappropriateness by first estimating the effect of potential inapropriateness on production and specialization and second using the model structure to translate these effects into country-level revenue productivity, $\Xi(\ell)$.

# 3. Background and Measurement: Agricultural Pests and Pathogens

To set-up our empirical analysis, we next provide background information about pest targeting in biotechnology. We then provide a detailed description of our main data source and measure of inappropriateness based on the dissimilarity of pest and pathogen environments for growing specific crops across different locations.

## 3.1 Pathogen Threats and Plant Breeding

Crop pests and pathogens (CPPs), which include viruses, bacteria, fungi, insects, and parasitic plants, are a dominant threat to agricultural productivity. Experts estimate that between 50-80% of global output is lost each year to CPP damage (Oerke and Dehne, 2004), which represents "possibly the greatest threat to productivity" across all environments (Reynolds and Borlaug, 2006, p. 3). In Brazil, a major agricultural producer, it is estimated that 38% of annual production is lost due only to insects (Gallo et al., 1988), amounting to \$2.2 billion in lost output per year (Bento, 1999). Prior to the development of transgenic corn, the Western Corn Rootworm alone caused \$1 billion in annual losses in the US and substantially more around the world (Gray et al., 2009). A critical focus of crop breeding, as a result, is developing resistance to damaging CPPs.

The most fundamental technique for breeding favorable plant traits, including those that confer CPP resistance, is mass selection: saving the seeds of the "best" plants from a given crop cycle, replanting them the next year, and repeating the process (McMullen, 1987, p. 41). This process naturally selects crop lineages with sufficient resistance to the local CPP environment. But it creates no selective pressure for resistance to non-present CPP threats, and such resistance is extremely unlikely to arise by chance mutation.

Historians have written extensively about how the environmental-specificity of traditional breeding severely limited the diffusion of agricultural technology in the 20th century. Moseman (1970, p. 71) argues that US programs during the 1960s to increase agricultural productivity in other countries via technological diffusion largely failed because of the "unsuitability of U.S. temperate zone materials [...] to tropical agricultural conditions." In a review of agricultural technology diffusion, Ruttan and Hayami (1973, p. 122) state that "ecological variations [...] among countries inhibit the direct transfer of agricultural technology." Reynolds and Borlaug's (2006) detailed account of one uncommonly successful program of international crop diffusion, the CIMMYT wheat program, makes clear the time and resources required to overcome these obstacles with coordinated international breeding.[19]

More recently, genetic modification (GM) has been added to the crop development toolkit. The vast majority of modern GM technology has directly related to conferring resistance to specific pests and pathogens (Vanderplank, 2012; Van Esse et al., 2020). In principle, direct access to a plant's genetic code side-steps the slow process of natural selection in the field and consequent obstacles to breeding for non-local environments. But, in practice, GM technology has been used almost exclusively for solving the pathogen threats facing high-income countries, due to these countries' higher demand (Herrera-Estrella and Alvarez-Morales, 2001).

An illustrative case study of how modern plant varieties are "locally" targeted comes from Bt varieties, a large and celebrated class of genetically modified plants. Bt varieties are engineered to express crystalline proteins, cry-toxins, that are naturally produced *Bacillus thuringiensis* bacteria ("Bt") and destructive toward specific insect species. Cry toxins are insecticidal because they bind receptors on the epithelial lining of the intestine and prevent ion channel regulation. Due to the specificity of intestinal binding activity, cry toxins are highly insect-specific. This feature, while crucial for limiting the Bt varieties' broader ecological impact, makes their development highly targeted to specific pest threats. The main targets for early Bt corn varieties were the European maize borer and maize rootworm (Munkvold and Hellmich, 1999), major threats in the US and Western Europe.[20] In other

[19]The authors describe, as one example, how cooperation between CIMMYT laboratories and the Brazilian Institute of Agricultural Research (EMPRAPA) enabled the production of semi-dwarf wheat varieties adapted to Brazil's acidic soil and distinct CPP environment. This process involved more than a decade of intense coordination and the development of a novel "shuttle breeding" program to breed alternate generations of plants in different locations.

[20]$\delta$-endotoxins produced by Bt were originally identified as candidate toxins specifically because of their effectiveness against these particular pests (Bessin, 2019). Monsanto's Bt corn varieties, MON863 and MON810 were developed with $\delta$-endotoxins selected for their effectiveness against maize rootworm, uncommon among Cry proteins (Galitsky et al., 2001).

parts of the world with different CPP threats, however, frontier Bt maize is neither commonly used nor effective. For example, in South Africa there is widespread resistance to Bt maize and production damaged caused by the maize stalk borer, which does not exist in the US but is widespread in sub-Saharan Africa (Campagne et al., 2017). Disparities in the international appropriateness of GM technologies therefore emerge as a result of focus on "rich-world pests."[21]

### 3.2 Plant Pest and Pathogen Data: The Crop Protection Compendium (CPC)

While the aforementioned examples highlight specific and extreme instances of pest-specificity, it is unclear whether they are representative of general biases agricultural technology. Our analysis, unlike existing field tests of specific varieties, has the advantage of being able to estimate the average effect of CPP mismatch across all crops and countries and connect it with an economic model to determine its aggregate consequences. We now introduce the key data that allow us to directly measure CPP dissimilarities across all global crop-specific ecosystems.

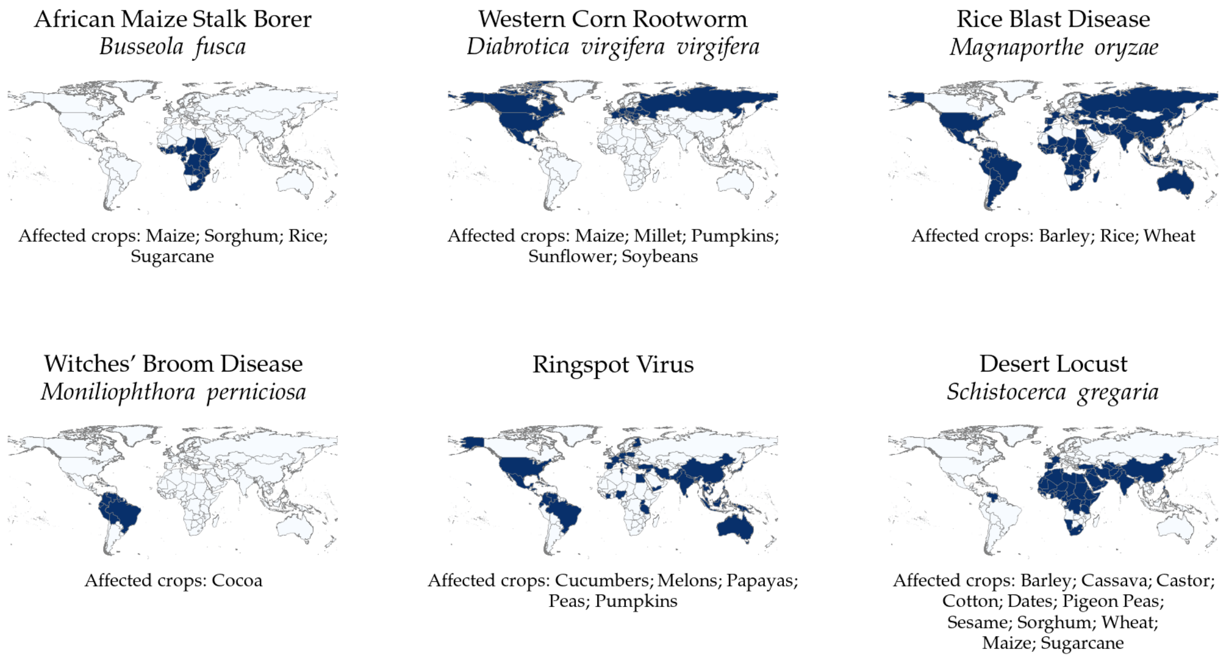
We source information on the global distribution of crop pests and pathogens from the Centre for Agriculture and Bioscience International's (CABI) Crop Protection Compendium (CPC). This database is the "world's most comprehensive site for information on crop pests," and provides detailed information on the geographic distribution and host species set for essentially all relevant plant pests and pathogens. Construction of the database began in the 1990s as a joint collaboration between CABI,the UN Food and Agriculture Organization, and the Technical Centre for Agricultural and Rural Cooperation (CTA). The goal of the project is to develop comprehensive and global coverage of crop diseases in order to better manage food production. The CPC was compiled through extensive searches of existing crop research, including the 460,000 research abstracts in the CABI database, as well as contributions from a range of governmental and international organizations, including the World Bank, the FAO, the United States Department of Agriculture (USDA), and the Consultative Group on International Agricultural Research (CGIAR) (Pasiecznik et al., 2005).[22]

In total, we compile information on 4,951 plant pests and pathogens, including viruses, bacteria, insects, fungi, and weeds. For each species, the CABI-CPC provides several key pieces of information. First, it provides information on the global geographic distribution. Figure 1 displays the distribution map for six pests, including the Maize Stalk Borer and Western Corn Rootworm, which were referenced in previous examples. For most countries, CABI reports whether the pest is present or not present in the country as a whole. For a handful of large countries—including Brazil and India,

[21] This pattern in GM development is not restricted to corn. The first varieties of Bt Cotton introduced in the early 1990s were focused on limiting the damage caused jointly by the tobacco budworm, cotton bollworm, and pink bollworm. In India, outbreaks of the pink bollworm in particular pose a major threat to cotton production (Fand et al., 2019). But frontier biotechnology has not adapted to patterns of Bt-resistance in India due to the lower relevance of the pink bollworm threat in the United States. Tabashnik and Carrière (2019) provide a review of pink bollworm resistance in global cotton populations.

[22] See here: `https://www.cabi.org/publishing-products/crop-protection-compendium/`. These data are the gold-standard for CPP measurement in population ecology and crop science; see, for example, studies by Bebber et al. (2013), Bebber et al. (2014a), Paini et al. (2016), and Savary et al. (2019).

Figure 1: Data on Example CPPs




*Notes*: Each map indicates CPP presence according to the CABI Crop Pest Compendium (CPC).

which we return to later—CABI reports state-level data on the presence of each CPP.[23]

Second, CABI reports all the host species that each pest or pathogen affects. For example, CABI reports that the African Maize Stalk Borer harms maize, sorghum, rice, and sugarcane, while the Western Corn Rootworm consumes maize, millet, pumpkins, sunflower, and soybeans, but not sorghum or sugarcane (Figure 1, top panel).[24] Our data contain information on 132 host species that are major crops, cross-referenced against the crops used in our subsequent analyses of biotechnology intellectual property and production.

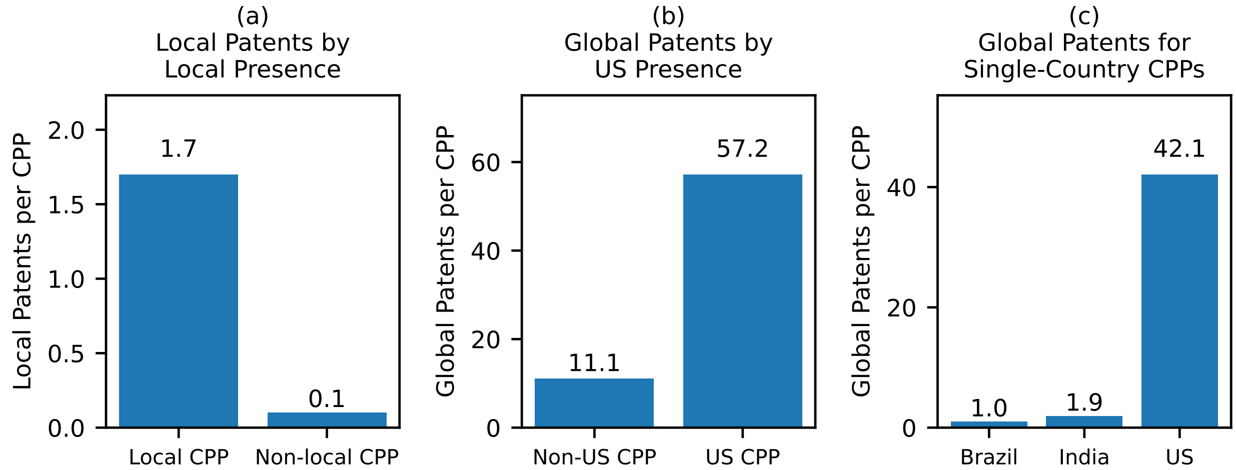
### 3.3 CPPs and the Direction of Global Innovation

With the CABI CPC data, it is possible to investigate empirically several features of global agricultural innovation discussed in Section 3.1 and built into our model. We identify all global biological or

[23]CABI also reports whether or not the pest or pathogen has been eradicated in a country, as well as whether each pest or pathogen is invasive or has high invasive potential. We use this information in our sensitivity analysis. The determinants of the cross-sectional global distribution of each pest and pathogen are not well understood by ecologists, and depend on "numerous [and] sometimes idiosyncratic" factors (Bebber et al., 2014b). Waage and Mumford (2008) and Shaw and Osborne (2011) come to a similar conclusion; while features of the environment, most prominently temperature and host prevalence, affect CPP presence, they often have limited predictive power and CPPs are often absent in ecologically habitable areas for reasons unrelated to human activity. Importantly also, Bebber et al. (2014b) find that CPP distributions measured from the CABI CPC appear unrelated to patterns of trade, travel, or tourism, suggesting that human activity plays a limited role in shaping the cross-sectional distribution of CPPs on average.

[24]We restrict attention to the host-pest relationships that are verified in the CABI database as opposed to those labeled as "data-mined" from articles and abstracts but not verified. This procedure retains 88% of all possible host-pest matches. Finally, note that the associations are *global* and not location-specific.

Figure 2: Global Patenting on CPPs




*Notes*: Graph (a) reports the average number of patented technologies developed in countries $\ell$ related to CPP threats $t$ if the CPP is present not present. Graph (b) reports the average number of patented technologies developed about CPPs that are not present in the US and CPPs that are present in the US. Graph (c) reports the number of patented technologies developed about CPPs that are present only in (i.e., endemic to) the countries specified on the $x$-axis.

chemical agricultural patents in the *PatSnap* database by searching for the scientific name of each CPP in all patent titles, abstracts, and descriptions.[25] We also identify the country of origin of each patent using *PatSnap*'s determination of the assignee's location. We document three facts about patenting at the country-by-CPP level, all consistent with the premise of the inappropriate technology hypothesis.

First, a large share of global innovation is focused on crop pests and pathogens (CPPs); 33% of all global biological and chemical agricultural patents are related at least one CPP in our sample.

Second, innovators focus substantially more on locally present CPPs. This pattern is apparent in the raw patent data: on average, over 17 times more patented technologies are developed for locally present CPPs compared to CPPs that are not present in the country of interest (panel (a) of Figure 2). We investigate this pattern more precisely by estimating the following regression:

$$y_{\ell,t} = \xi \cdot \text{Local CPP}(\ell, t) + \chi_\ell + \chi_t + \varepsilon_{\ell,t} \tag{3.1}$$

where the unit of observation is a CPP-year and Local CPP$(\ell, t)$ is an indicator that equals one if CPP $t$ is present in country $\ell$. $y_{\ell,t}$ is the number of patented technologies developed in country $\ell$ related to CPP threat $t$, transformed by the inverse hyperbolic sine, and $\chi_\ell$ and $\chi_t$ absorb country and CPP fixed effects. $\xi$ captures the extent to which innovation is disproportionately targeted toward local CPP threats. Table A1 reports our estimates. We estimate that $\xi > 0$ in Equation 3.1, and it remains

[25]The full set of biological or chemical agricultural patents are all those that comprise Cooperative Patent Classes (CPC) A01H and A01N. Individual patents can link to multiple CPPs if the patent references multiple species.

large and significant focusing on either the intensive or extensive margin separately (columns 2-3).

Third, substantially more technology is developed to combat CPPs that exist in high-income countries like the US. Panel (b) of Figure 2 demonstrates that CPPs present in the US have a more than five-fold higher quantity of patents on average than those not present in the US. Table A2 reports estimates from an augmented version of (3.1) in which Local CPP$(\ell, t)$ is interacted either with an indicator that equals one if $\ell$ is the US (columns 1-3) or (log of) per-capita GDP of $\ell$ (columns 4-6). The impact of a locally present CPP on innovation is substantially larger in high-income countries, consistent with greater overall R&D intensity. Finally, panel (c) of Figure 2 shows one particularly striking cut of the data: the number of patents about CPPs that are present *only* in, or endemic to, the US dwarfs the number for CPPs present only in two of the world's largest, but significantly less research intensive, agricultural economies, Brazil and India.

This analysis, taken together, documents that (i) a large share of global agricultural innovation is focused on CPPs and (ii) much of this research is highly localized. The end result is a far greater focus on CPP threats present in high-income, research-intensive countries. These findings are consistent with the set-up of our model of endogenous technology.

### 3.4 Measuring Inappropriateness: CPP Mismatch

The remainder of our empirical analysis starts from the premise of unequal research intensity and studies how ecological differences affect technology diffusion and production. In the model, the scalar summary of ecological difference was the measure of non-common ecological features or CPP threats, $\delta(k, \ell, \ell')$. In the data, using our lists of locally present CPPs affecting crop $k$ in each location $\ell$ or $\ell'$, we compute the following measure of "CPP Mismatch" at the location-pair-by-crop level which is the same up to a necessary normalization:

$$\text{CPP Mismatch}_{k,\ell,\ell'} = 1 - \frac{\text{Number of Common CPPs}_{k,\ell,\ell'}}{\left(\text{Number of CPPs}_{k,\ell} \times \text{Number of CPPs}_{k,\ell'}\right)^{1/2}} \tag{3.2}$$

The measure, which has the form of one minus a correlation or cosine similarity, equals zero when $\ell$ and $\ell'$ have all the same CPPs for crop $k$ and equals one when $\ell$ and $\ell'$ have no CPPs in common for crop $k$. In the language of ecology, as discussed in a review chapter on biological similarity by Jost et al. (2011), our CPP mismatch formulation in (3.2) is one of several standard divergence (one-minus-similarity) measures that satisfy basic properties of *density invariance*, *replication invariance*, and *monotonicity*. Heuristically, this means that the divergence or similarity measures provide consistent results regardless of the total number of species or population of any individual species in $\ell$ or $\ell'$.[26]

[26] We will also, as a robustness check throughout our analysis, supplement our main measure with the simplest and most historical measure of divergence due to Jaccard (1900, 1901) which counts the fraction of non-shared species:

$$\text{CPP Mismatch}^{J}_{k,\ell,\ell'} = 1 - \frac{\text{Number of Common CPPs}_{k,\ell,\ell'}}{\text{Number of Unique CPPs}_{k,\ell\cup\ell'}} \tag{3.3}$$

CPP Mismatch varies at both the country-pair level, fixing crops, and the crop level, fixing country pairs. The *country-level variation* is illustrated by Figure 1: different countries are endowed with different CPPs. The *crop-level variation* is due to the fact that each CPP only affects a particular set of crops: for each example in Figure 1, the set of affected crops varies substantially. Depending on the identities of each country's locally present CPPs, a single pair of countries will have different CPP distances across crops. To give one example of this variation, Appendix Figure A1 shows the histogram of all countries' CPP mismatch with the US for wheat and sugarcane and identifies the observations for Brazil and India. For wheat, India is very slightly more similar to the US than Brazil is. For sugarcane, Brazil is substantially more similar to the US than India is. Having these two sources of variation allows us to fully control for any differences across countries or crops in our empirical analysis.

Our baseline measure of CPP mismatch uses all CPPs in the CABI database in order to capture the full extent of CPP differences around the world today. To investigate the potential role of invasive species, which are an important mechanism but also potentially endogenous to human behavior, we use the CABI Invasive Species Compendium (ISC) to identify all invasive and high-invasive-potential CPPs and drop them from the calculation of CPP Mismatch. The ISC data and corresponding analysis are discussed in more detail in Appendix Section B.1.

We also investigate the importance of non-CPP differences in ecology and geography—including temperature, precipitation, and soil characteristics—as additional shifters of appropriateness. Appendix Section B.2 discusses our measurement of alternative sources of crop-by-country-pair geographic mismatch, as well as all empirical results using these alternative measures alongside our baseline CPP mismatch measure. In summary, we find that differences in other agro-climatic features also inhibit technology transfer and distort specialization; that these effects are independent from the effects of CPP mismatch; and that the effects of CPP mismatch are larger. These results, along with the anecdotal evidence about plant breeding and technology diffusion from earlier in this section, motivate our focus on CPP differences in the primary analysis.

## 4. Main Results: Technology Diffusion

In this section, we investigate the relationship between inappropriateness and technology diffusion. Our empirical strategy uses variation in inappropriateness and technology transfer at the country-pair-by-crop level, combining our CPP Mismatch measure introduced in the previous section with a new database of the invention and international transfer of plant varieties.

This metric has the same range (0 to 1) and interpretation of extreme values as our baseline, but different properties for intermediate levels of similarity.

### 4.1 Data: The UPOV Plant Variety Database

We measure the development and international transfer of biotechnology inventions using a novel dataset of all global instances of intellectual property protection for crop varieties. We obtained these data from The International Union for the Protection of New Varieties of Plants (UPOV), the inter-governmental organization tasked with designing, promoting, and administering systems of intellectual property protection for plant varieties around the world.[27] The data provide comprehensive coverage of all plant variety certificates, an internationally standardized form of intellectual property, across the member countries identified in the map in Figure A2.[28]

In order to obtain protection under UPOV, a variety must be new, distinct, uniform (i.e. identical across plants within a generation), and stable (i.e. identical across generations); this means that each variety in the database represents a unique technology that is usable in production.[29] Moreover, this set of variety characteristics is relatively straightforward to document, meaning that barriers to obtaining protection—both in terms of legal fees and the burden associated with documenting the inventive step—are limited. The ease of obtaining protection helps ensure that the UPOV database captures a large share of varieties in circulation.[30] Finally, a breeder must protect a variety separately in each country where they want legal enforcement, meaning that observing that a variety is protected in a particular country is a strong indication that the variety was transferred to that country, and the absence of protection is a strong indication that the variety was not transferred to that country.[31]

For each certificate, we observe (i) the date of issuance; (ii) the country of issuance; (iii) the plant species; and (iv) a unique "denomination" identifier associated with the variety. The UPOV Convention of 1991 stipulates that the denomination of a specific plant variety must be consistent across member countries.[32] That is, wherever in the world a denomination code is observed in the database, it corresponds to a single, unique plant variety. This allows us to track the diffusion of individual varieties, which we also refer to interchangeably as "technologies," across countries. The certificate data, when cross-linked to a list of major agricultural crops and screened for duplicate entries, consists of 458,034 total variety registrations, spanning 62 countries, 109 crops, and 236,529 unique denominations.

Figure 3 displays a snapshot of the raw UPOV data. These five rows are from the section of the database on cotton varieties registered between 1999 and 2003. This example consists of three unique denominations (Sicot 41, Sicot 53, and Sicot 71) registered across three countries (Australia, Argentina,

[27]Our project required a formal application process and approval from the UPOV Council.

[28]This set notably excludes several large agricultural producers in South Asia, North Africa, and Sub-Saharan Africa, on account of these countries' imperfect recognition of plant variety intellectual property. We return to this topic at various points in the analysis, including with an alternate measure of variety presence in Sub-Saharan Africa (see Section B.4).

[29]For more detail, see the description here: `https://www.upov.int/overview/en/conditions.html`.

[30]This helps ameliorate concerns associated with measuring technology using patent data, which is often skewed toward large, private sector firms due to the high financial barriers to obtaining protection.

[31]For additional detail, see here: `https://www.upov.int/about/en/faq.html#QG90`.

[32]This stipulation is described in Article 20.5 ("Same denomination in all Contracting Parties") of the most recent (1991) revision of the UPOV Convention (Union for the Protection of New Varieties of Plants, 1991). Further clarification is provided in the "Explanatory Notes" on variety denominations (Union for the Protection of New Varieties of Plants, 2015).

Figure 3: Example Rows from UPOV Data Set

| UPOV Code | Country | Denomination | Botanical Name | Common Name | App. Date |
|---|---|---|---|---|---|
| *GOSSY_HIR* | AU | Sicot 53 | Gossypium hirsutum | Cotton | 14-Sep-99 |
| *GOSSY_HIR* | AU | Sicot 41 | Gossypium hirsutum | Cotton | 14-Sep-99 |
| *GOSSY_HIR* | AR | Sicot 41 | Gossypium hirsutum L. | Algodonero | 13-Aug-01 |
| *GOSSY_HIR* | AU | Sicot 71 | Gossypium hirsutum | Cotton | 07-Aug-02 |
| *GOSSY_HIR* | BR | Sicot 53 | Gossypium hirsutum L. | Algodao | 11-Nov-03 |

*Notes*: This figure reports example rows from the UPOV PLUTO database. The rows reported are those related to unique varieties Sicot 53, Sicot 41, and Sicot 71, developed by Australia's Commonwealth Scientific and Industrial Research Organization. The UPOV Denomination Code uniquely identifies specific varieties wherever they appear in the world.

and Brazil). The data reveal that Sicot 53 cotton was first registered in Australia in 1999 and later in Brazil in 2003. Sicot 41 cotton was also introduced in Australia in 1999 and transferred to Argentina in 2001. Finally, Sicot 71 cotton was introduced in Australia in 2002, but was never introduced in any other country.[33]

We generalize the above example into a method for tracking the diffusion of specifically identified pieces of technology, like Sicot 41 cotton, between locations. For every unique denomination in the data, we identify a country of first appearance. We use the country of first appearance as the origin country since this is most likely to be the market for which the variety was first developed.[34] We then count, in any given time period, the number of varieties identified for a crop $k$, newly registered in country $\ell$, and originating from country $\ell'$. This will be our primary measure of technology diffusion between country pairs at the crop level. For our main analysis, we focus on a static cross section and sum over all final registration events after 2000.[35]

Echoing the previous discussion about the concentration of innovation in richer countries, 67% of all recorded varieties are first reported in one of the United States, Canada, or a European Union member state.[36] Among all varieties, 34% are transferred at least once between countries. This number increases to 49% when sub-setting to varieties first reported in the aforementioned set of countries, offering a first indication that varieties from "leader countries" are more often spread worldwide. Figure A3 presents summary statistics on the likelihood of variety transfer in our sample and visualizes the network structure of variety transfers across countries. Appendix B.3 also presents a

[33]Sicot cotton is a product of Australia's Commonwealth Scientific and Industrial Research Organization, an Australian governmental agency, which incorporates genetic modification to achieve "desired fibre quality, disease resistance and yield." See here: `https://csiropedia.csiro.au/cotton-breeding-and-new-cotton-varieties/`.

[34]This avoids potential issues associated with using the country of the innovating firm or firm headquarters. For example, while Monsanto was headquartered in the US during our sample period, is invested substantially in developing soybean technology tailored to the Brazilian market. Our strategy would correctly identify the intended beneficiary of this technology as Brazil, rather than the US.

[35]Note that we do not truncate the data to post-2000 when identifying country of origin, so a variety like Sicot 41 in the example (first registered in 1999 in Australia) is still in our final data set as a variety transferred to Argentina in 2001.

[36]These constitute 26 of the 62 countries in our sample.

more detailed analysis of the global direction of innovation in the UPOV variety database, mirroring our analysis of CPP-level patents in Section 3.3. There is a strong concentration of innovation in crops cultivated in high-income and in crops cultivated in countries that enforce intellectual property protection for plant biotechnology, and that this effect is driven by substantial home bias toward locally abundant crops.

### 4.2 Empirical Model

Our main estimating equation is the following linear regression, which is the empirical analog of Equation 2.4 in Proposition 1:

$$y_{k,\ell',\ell} = \beta \cdot \text{CPP Mismatch}_{k,\ell',\ell} + \chi_{\ell,\ell'} + \chi_{k,\ell} + \chi_{k,\ell'} + \varepsilon_{k,\ell,\ell'} \tag{4.1}$$

where $k$ indexes crops, $\ell$ indexes technology receiving countries, and $\ell'$ indexes technology sending countries. The outcome $y_{k,\ell',\ell}$ is a monotone transformation of the number of unique varieties of crop $k$ developed in $\ell'$ and transferred to $\ell$ between 2000-2018. Since there are many zeroes in the varieties data, we report the effect separately for the intensive margin with log biotechnology transfers, the extensive margin with an indicator for any transfer, and the inverse hyperbolic sine (asinh) transformation which blends the two margins. Our baseline specification includes all possible two-way fixed effects: origin-by-destination fixed effects, crop-by-origin fixed effects, and crop-by-destination fixed effects. These absorb, for example, the fact that certain countries persistently demand or develop more technology for particular crops, as well as any crop-invariant features of country pairs (e.g. physical and cultural distance, common geography, trade linkages, etc.).[37] Standard errors are double-clustered by origin and destination.

The main hypothesis is that $\beta < 0$, which would indicate that the local focus and context specificity of innovation depresses technology diffusion and that, on average, biotechnology flows less when technology is inappropriate. We may find no effect, however, if the context-specific component of technological progress or local research spillovers are relatively small, or if technology diffusion is sufficiently "inelastic" with respect to incentives.

While estimates of $\beta$ from Equation 4.1 capture the average relationship between CPP mismatch and technology transfer, Proposition 1 demonstrated that the effect could be very different across crop-origin pairs; in particular, the marginal effect of ecological dissimilarity should be larger when the sending country is very active in research for crop $k$. To empirically investigate this idea, we also estimate versions of the following augmented version of (4.1) that parameterizes heterogeneity in the main effect :

$$y_{k,\ell',\ell} = \beta_1 \cdot \text{CPP Mismatch}_{k,\ell',\ell} + \beta_2 \cdot F_{k,\ell'} \times \text{CPP Mismatch}_{k,\ell',\ell} + \chi_{\ell,\ell'} + \chi_{k,\ell} + \chi_{k,\ell'} + \varepsilon_{k,\ell,\ell'} \tag{4.2}$$

[37] The exact interpretation of these effects is described in Proposition 1 and its proof.

Table 1: CPP Mismatch Inhibits International Technology Transfer

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | Biotech Transfer (asinh) | Any Biotech Transfer (0/1) | log Biotech Transfer |
| CPP Mismatch (0-1) | -0.0624** | -0.0275** | -1.202*** |
| | (0.0235) | (0.0106) | (0.386) |
| Crop-by-Origin Fixed Effects | Yes | Yes | Yes |
| Crop-by-Destination Fixed Effects | Yes | Yes | Yes |
| Origin-by-Destination Fixed Effects | Yes | Yes | Yes |
| Observations | 204,287 | 204,287 | 5,791 |
| R-squared | 0.439 | 0.383 | 0.797 |

*Notes:* The unit of observation is a crop-origin-destination. All possible two-way fixed effects are included in all specifications. The dependent variable is listed at the top of each column. Standard errors are double-clustered by origin and destination and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

where $F_{k,\ell'}$ is an indicator variable that equals one for the countries $\ell'$ that we identify as the biotechnological leaders for crop $k$. We have two strategies for defining $F_{k,\ell'}$. The first is to treat the United States as the frontier for all crops, or set $F_{k,\ell'} = \mathbb{I}[\ell' = \text{US}]$. This method is motivated by the United States' pre-eminence in modern agricultural research.[38] The second is to identify a set of crop-specific "leaders" $T_N(k)$ in the UPOV data, based on being among the top $N$ countries in variety registrations for $k$.[39] This data-driven approach sets $F_{k,\ell'} = \mathbb{I}[\ell' \in T_N(k)]$, and is parameterized by the list length $N$. In this specification, $\beta_2 < 0$ captures the difference in the marginal effect of inappropriateness on technology diffusion when the origin country is a leader in biotechnology development.

### 4.3 Results

Estimates of Equation 4.1 are reported in Table 1. On all margins, we find that CPP mismatch significantly inhibits the international flow of technology. The intensive-margin estimates from column 3 imply that CPP mismatch inhibits 30% of international technology transfer for the median crop and country-pair, suggesting that even in the full sample crops and countries CPP mismatch and the inappropriateness of foreign technology is a major barrier to international technology diffusion.

Before proceeding, we probe the sensitivity of the baseline estimates. We first reproduce our results under different measurement strategies for ecological differences. Column 1 of Table A3

[38] The US alone produced 30% of citation-weighted global agricultural science publications. The US is also the global leader in patented agricultural technology and produces three times as many patents as the next highest country (Japan). 52% of agricultural research and development companies are incorporated in North America and US inventors generate roughly 1.5 thousand patents for plant modification and 1 thousand patents for cultivar development *per year* (Fuglie, 2016).

[39] By counting registrations, we multiply count unique denominations that are registered in multiple countries. This is intentional, to capture the countries whose technologies are most diffusive. Similar results are obtained by doing the analysis at the denomination level.

reproduces our baseline estimates for reference. In column 2 , we show our results are stable using the Jaccard (1900, 1901) mismatch metric (see Equation 3.3). In column 3, we show the same using an alternative CPP mismatch classification that counts CPPs as "present" if CABI lists any information about them, including whether they have been eradicated in the past.[40] In Appendix B.1, we discuss how we can use the CABI data to identify possible species invasions in recent history, which could be affected by crop-level trade or connectedness between countries, and show the stability of our results to excluding all invasive CPPs. Thus, the findings are not driven by CPP eradications or invasions, both of which are rare compared to the full set of global CPP threats.

We also explore whether the results are influenced by links across countries that are not related to differences in the CPP environment. All specifications include origin-by-destination fixed effects, so any relevant omitted variable must also vary *across crops* within a country pair. Features like geographic or cultural distance between countries are fully absorbed by the country pair fixed effects. In column 4 of Table A3, we control for an indicator that equals one if countries $\ell$ and $\ell'$ engage in bilateral final good trade for crop $k$. In column 5, we control for (log of) the geographic distance between all country pairs interacted with a full set of crop fixed effects, allowing the effect of distance to vary flexibly across crops (for instance, via crop-specific trade costs). In columns 6 and 7, we exclude from the sample origin-destination pairs within 1000km or 2000km of each other respectively. Each exercise produces stable results. Finally, Table A13 reports results after controlling for several non-CPP measures of ecological dissimilarity across crops and country-pairs, and again the estimates are very similar. This analysis is discussed in detail in Appendix B.2.

We next identify the effect of ecological mismatch *relative to the frontier* on technology diffusion. Table 2 reports estimates of (4.2), which includes an interaction term between CPP mismatch and an indicator that equals one if the origin country is a frontier technology developer. The blended, extensive, and intensive margin effects are reported in Panels A, B, and C, respectively; and our definitions of the frontier as the US, $T_1(k)$, $T_2(k)$, and $T_3(k)$ are used in columns 1-4. In the extensive and blended-margin specifications, we find strong, significant evidence of $\beta_2 < 0$; in the intensive-margin specification, we have consistently negative point-estimates, which are statistically significant in one of four cases. The effect of CPP mismatch on technology diffusion is considerably larger for research intensive origins, and in some specifications we find that CPP mismatch with countries outside of the frontier has not impact on technology diffusion. For example, in columns 3-4 of Panel A, the marginal effect of CPP distance on (asinh) technology diffusion is roughly thirty times larger for frontier origin markets and statistically indistinguishable from zero for non-frontier origin markets.

These estimates imply that high ecological dissimilarity to the frontier can leave a country with little or no appropriate modern technology. Interpreted via the model, they are consistent with a large context-specific component of modern technology and local research spillovers in frontier countries. As a result, ecological mismatch substantially reduces the cross-border transfer of biotechnology.

[40]Potential eradications are quite rare. The number of CPP-country-crop triads increases by under 3% when using the "broad" CPP presence classification.

Table 2: CPP Mismatch with Frontier Countries and Technology Transfer

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Frontier defined as: | United States | Top Variety Developer | Top 2 Variety Developers | Top 3 Variety Developers |
| | *Panel A: Dependent Variable is (asinh) Biotech Transfers* | | | |
| CPP Mismatch (0-1) | -0.0571** | -0.0453** | -0.0330 | -0.0207 |
| | (0.0216) | (0.0215) | (0.0199) | (0.0196) |
| CPP Mismatch (0-1) x Frontier (0/1) | -0.392*** | -1.237*** | -1.076*** | -1.076*** |
| | (0.0313) | (0.290) | (0.249) | (0.249) |
| Observations | 204,287 | 204,287 | 204,287 | 204,287 |
| R-squared | 0.439 | 0.442 | 0.444 | 0.444 |
| | *Panel B: Dependent Variable is Any Biotech Transfer (0/1)* | | | |
| CPP Mismatch (0-1) | -0.0241** | -0.0229** | -0.0181* | -0.0136 |
| | (0.00956) | (0.00986) | (0.00917) | (0.00884) |
| CPP Mismatch (0-1) x Frontier (0/1) | -0.254*** | -0.332*** | -0.343*** | -0.322*** |
| | (0.0142) | (0.0699) | (0.0623) | (0.0535) |
| Observations | 204,287 | 204,287 | 204,287 | 204,287 |
| R-squared | 0.383 | 0.384 | 0.385 | 0.385 |
| | *Panel C: Dependent Variable is log Biotech Transfers* | | | |
| CPP Mismatch (0-1) | -1.161*** | -1.084*** | -1.154*** | -0.852** |
| | (0.364) | (0.350) | (0.322) | (0.381) |
| CPP Mismatch (0-1) x Frontier (0/1) | -0.698 | -0.694 | -0.173 | -0.892** |
| | (1.248) | (0.423) | (0.503) | (0.437) |
| Observations | 5,791 | 5,791 | 5,791 | 5,791 |
| R-squared | 0.797 | 0.797 | 0.797 | 0.797 |
| Crop-by-Origin Fixed Effects | Yes | Yes | Yes | Yes |
| Crop-by-Destination Fixed Effects | Yes | Yes | Yes | Yes |
| Country Pair Fixed Effects | Yes | Yes | Yes | Yes |

*Notes:* The unit of observation is a crop-origin-destination. The definition of a leader in each specification is noted at the top of each column and the dependent variable is noted in the panel heading. Standard errors are double-clustered by origin and destination and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

# 5. Main Results: Production

The previous section established that potential inappropriateness determined by ecological mismatch inhibits technology transfer. We now study how ecological differences relative to frontier innovators affect global production and specialization.

## 5.1 Data and Measurement

### 5.1.1 Agricultural Production

We compile data on crop output, trade (imports and exports), and prices from the UN Food and Agriculture Organization statistics database (FAOSTAT). We also compile sub-national agricultural output data from the latest nationally representative agricultural census for both Brazil and India.

The Brazilian data are from the 2017 round of the Censo Agropecuario, and they cover 49 crops. The Indian data are from the ICRSAT Database and constructed from the 2015 Agricultural census, and they cover 20 states and 20 crops.[41]

### 5.1.2 Mismatch with the Frontier

Mapping our analysis to the predictions of Proposition 2 requires taking a stand on "which inappropriateness matters" for determining a given country's production, or from where that country sources its technology. Since we lack detailed data on the country of origin for the crop-specific inputs used in each market, we instead use two more heuristic but parsimonious strategies to measure each country's ecological dissimilarity to the *frontier technology producers*, as introduced in Section 4.2.

The first, and simpler, strategy is to assume that the United States produces the frontier technology for all crops and set $\text{CPPMismatchFrontier}^{\text{US}}_{k,\ell} = \text{CPPMismatch}_{k,\ell,\text{US}}$. In the model, this method is exactly correct if the United States were the sole producer of technology. In reality, nearly fifty percent of private research investment takes place in the US, representing a large share of global innovation (Fuglie, 2016). Our second strategy is to define the technological frontier for each crop based on the frequency of variety releases in the UPOV data. Given a set $T_N(k)$ of the $N$ top countries for $k$-variety releases, we calculate:

$$\text{CPPMismatchFrontier}^{\text{Est}}_{k,\ell} = \sum_{\ell' \in T(k)} \left(\text{Share Varieties}^{\text{UPOV}}_{k\ell'}\right) \times \left(\text{CPP Mismatch}_{k,\ell,\ell'}\right) \tag{5.1}$$

where $\text{CPP Mismatch}_{k,\ell,\ell'}$ is our main bilateral measure defined in Equation 3.2. This method picks up geographic variation in technological leadership, but relies on cross-national comparisons of variety release intensity.[42] For our baseline results, we use $N = 2$; however, the results are very similar for alternative values for $N$.

These strategies for defining frontier innovators are further motivated by the results in Table 2, showing that CPP mismatch with the US or countries in $T(k)$ have a disproportionate negative effect on biotechnology diffusion. In fact, in some specifications, CPP mismatch with countries *outside* this set of frontier countries has zero effect on technology diffusion (e.g. columns 3-4 of Panel A).

In practice, the multiple measures of CPPMismatchFrontier have a similar distribution across crops and space and a strong positive correlation with one another. In a univariate regression of the former on the latter, the coefficient 0.93 (0.047) and $R^2$ is 0.91. The underlying reason is that our identified technological leaders, in the majority of cases, are subsets of the US, Canada, and temperate countries in Western Europe. This foreshadows the fact that our main findings are similar using either measure.

[41] For a description of the ICRSAT data, see here: `http://data.icrisat.org/dld/src/about-dld.html.`

[42] In the model, this can be mapped to case in which only the countries $\ell \in T(k)$ produce technology for $k$, productivity $\Theta(k, \ell)$ is linearly approximated around a steady state with $\delta(k, \ell, \ell') \equiv 0$ for all $\ell' \in \Theta(k, \ell)$, and $\text{ShareVarieties}_{k\ell'}$ equals the fraction of farms that would use $\ell'$ technology if all technology were equally appropriate.

#### 5.1.3 Direct Effects of the Local Environment

In the model, the relationship between ecological mismatch and production was correctly specified conditional on measurements of the parameter $\omega(k,\ell)$, local innate suitability for growing crop $k$ in country $\ell$ (see Proposition 2). To directly capture the impact of local suitability on output in our analysis, we use two measurement strategies. First, we directly measure crop-specific production as predicted by local geography from the FAO Global Agro-Ecological Zones (GAEZ) model and database (see, e.g., Costinot and Donaldson, 2012; Costinot et al., 2016). We compute total predicted production under GAEZ's low-input, rain-fed scenario, which holds fixed background differences in input use and technology, on land area within a country on which a given crop was grown according to a cross-section in 2000, as measured by the *EarthStat* database of Monfreda et al. (2008). While this method parsimoniously summarizes agronomic predictions of innate suitability, it is only available for 34 of our 132 crops.

Our second approach is to compile a larger set of environmental variables and then use post-double LASSO (Belloni et al., 2014) to select an appropriate set of control variables, tantamount to specifying our own crop-specific empirical models for suitability. We first construct fixed effects for the 200 "most geographically prevalent" CPPs, as determined by the number of countries in which they are present, and the 200 "most agriculturally prevalent" CPPs, as determined by the number of host species that they infect. We also construct measures of average temperature, precipitation, elevation, ruggedness, the growing season, and soil acidity, clay content, silt content, coarse fragment content, and water capacity at the crop-by-country level, by averaging these variables over the historical planting locations from the *EarthStat* database. Appendix B.2 describes these data in more detail.

### 5.2 Empirical Model

We estimate the following model which is the empirical analog of Equation 2.7 in Proposition 2:

$$y_{k,\ell} = \beta \cdot \text{CPPMismatchFrontier}_{k,\ell} + \chi_\ell + \chi_k + \Omega'_{k\ell}\Gamma + \varepsilon_{k,\ell} \tag{5.2}$$

The outcome $y_{k,\ell}$ is average production from 2000 to 2018 in log physical units. All specifications include country and crop fixed effects ($\chi_\ell$ and $\chi_k$), which capture any aggregate differences across countries (e.g., income, productivity) or crops (e.g., market size, price). Depending on the specification, we include a subset of proxies for innate suitability in the vector $\Omega_{k,\ell}$. The coefficient of interest is $\beta$, which captures the effect of CPP dissimilarity from technology producing countries on features of agricultural production.

### 5.3 Results

Our baseline estimates of (5.2) are reported in Table 3. In columns 1-4, CPP mismatch with the frontier is parameterized as crop-specific mismatch with the US and, in columns 5-8, it is parameterized as

Table 3: CPP Mismatch Reduces Agricultural Output

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Dependent Variable is log Output | | | | | | | |
| | CPP Mismatch with the US | | | | CPP Mismatch with the Estimated Frontier | | | |
| CPP Mismatch (0-1) | -9.285*** | -10.60*** | -9.325*** | -8.454*** | -7.136*** | -5.721*** | -7.202*** | -6.288*** |
| | (1.199) | (3.024) | (0.617) | (0.652) | (0.959) | (0.663) | (0.461) | (0.501) |
| log(FAO-GAEZ-Predicted Output) | | 0.298*** | | | | 0.353*** | | |
| | | (0.0814) | | | | (0.0499) | | |
| *Included in LASSO Pool:* | | | | | | | | |
| Top CPP Fixed Effects | - | - | Yes | Yes | - | - | Yes | Yes |
| Ecological Features x Crop Fixed Effects | - | - | No | Yes | - | - | No | Yes |
| Controls in LASSO Pool | - | - | 335 | 3935 | | | 335 | 3935 |
| Crop Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6,926 | 2,353 | 6,931 | 6,069 | 6,704 | 2,353 | 6,707 | 5,903 |
| R-squared | 0.599 | 0.617 | | | 0.600 | 0.609 | | |

*Notes:* The unit of observation is a country-crop pair. Columns 1-4 use CPP mismatch with the US and columns 5-8 use CPP mismatch with the estimated set of technological leader countries. Columns 1-2 and 5-6 report OLS estimates and columns 3-4 and 7-8 report post double LASSO estimates. Country and crop fixed effects are included in all specifications, and included in the amelioration set in thet post-double LASSO specifications. The Top CPPs are defined as the top 200 CPPs defined by (i) the number of countries in whch they are present and (ii) the number of host crops that they infect. Since the two sets overlap, the total number is 335. The set of ecological features includes: temperature, precipitation, elevation, ruggedness, growng season days, soil acidity, soil clay content, soil silt content, soil coarse fragment volume, and soil water capacity. Standard errors are double-clustered by crop and country and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

crop-specific mismatch with the crop-specific estimated frontier country set, $T_2(k)$. In both cases, we estimate a large and significant negative coefficient. Our estimates imply that a one standard deviation in increase in CPP mismatch with frontier countries lowers output by 0.51 standard deviations.

The specifications in columns 1 and 5 only include the CPP mismatch measure, along with crop and country fixed effects, on the right-hand side of the regression. The remaining columns show the stability of these estimates under each of our control strategies for innate suitability. In columns 2 and 6, we include the FAO GAEZ agronomic model-derived output estimate as a control. In columns 3 and 7, we show estimates from the post-double LASSO control strategy using the top CPP fixed effects. In columns 4 and 8, we expand the LASSO pool to include the full set of country-level geographic covariates, and their square, interacted with crop-fixed effects, to allow for crop-specific effects of each characteristic.[43] Results are stable in each variation of the control strategy.

Table A4 reports an analogous set of estimates to Table 3 with log of area harvested (instead of output) as the dependent variable. Consistent with the predictions of the Fréchet model for selection effects, we find statistically indistinguishable magnitudes relative to our main estimates for production. Economically, this implies that agricultural allocations eliminate cross-crop differences in marginal products. As we will discuss extensively in Section 7, we can use the model structure plus the measured effects on production to infer overall productivity effects consistent with the model.

[43]Thus, all control vary at the country-by-crop level. When we include all aforementioned controls, the LASSO pool contains 3.935 potential covariates. Throughout, we include crop and country fixed effects in the LASSO amelioration set.

Table A5 investigates the impact of CPP mismatch on additional features of agricultural production and output. First, we document that CPP mismatch with the frontier is significantly negatively correlated with crop-specific exports (column 2), and positively (albeit insignificantly) correlated with crop-specific imports (column 3). Second, we document that CPP mismatch is significantly positively correlated with producer price volatility. This finding indicates that the appropriateness of frontier technology might not only raise average productivity but also increase producers' ability to withstand periodic negative productivity shocks.[44] The negative relationship with producer price volatility is similar even after holding total output fixed (columns 5 and 7).

The stability of all findings after accounting for local suitability is consistent with the fact that, *ex ante*, there is no reason to expect that the locations with the best biotechnology firms for producing seeds for a particular crop are also innately the best places for growing that crop. Thus, there is no reason to believe that being ecologically "distant" from technology producing countries is tantamount to being ecologically "bad." Indeed, in the US there is a long history of science and technology development to confront crop disease and the challenging pathogen environment (Olmstead and Rhode, 2008).[45] Consistent with this history of ecological challenges in what would become a highly agriculturally productive country, existing empirical evidence suggests that variation in local land suitability plays a limited role in explaining global productivity differences (Adamopoulos and Restuccia, 2018).

Our results, on the other hand, suggest that the indirect role of geography via production technology, or the endogenous determination of "good geographies" that resemble technological leaders', is an important determinant of production patterns. To make this point explicitly, Appendix Section B.5 documents that the unprecedented rise of US biotechnology research since the 1990s is associated with shifts in global specialization toward crops and countries where US technology is more appropriate. In particular, we show that CPP mismatch with the US is negatively associated with *changes* in crop-by-country level output since 1990, and that the same is not true for CPP mismatch with Europe, where biotechnology research grew substantially less during the past two decades. These results, along with related estimates investigating the changing locations of breeding during the Green Revolution which we turn to in Section 6.1, further indicate that our findings are not driven by a static omitted variable, and that "good geographies" change with the focus of innovation.

## 5.4 Sensitivity Checks

### 5.4.1 Additional Controls and Measurement

The results in Table 3 are very similar after including a range of additional controls. Table A6 documents that the results are very similar including crop-by-continent fixed effects, which allow us to focus on even more geographically precise variation in the inappropriateness instrument. Table A7

[44] Bad insect outbreaks are a commonly cited example. See Stone (2020) for a discussion of recent locust outbreaks in East Africa and their economic impact.

[45] In fact, during its early history, the US government made a major effort to recover plant varieties from around the world in order to increase farm productivity and promote agricultural resilience (Kloppenburg, 2005).

shows that results are similar after controlling for a broad spectrum of country-level characteristics, all interacted with crop fixed effects, which rules out confoundedness with crop-specific effects of other determinants of income.[46] The results are also similar after purging the CPP mismatch measure of variation driven by invasive species (Appendix B.1) and accounting for mismatch with the frontier in non-CPP ecological characteristics (Appendix B.2). Inappropriateness measured using non-CPP ecological characteristics also depress technology diffusion also reduces output; however, this effect is independent from and smaller than the effect of CPP distance (Table A14).

### 5.4.2 Falsification Tests

If our main estimates capture the impact of inappropriateness on technology diffusion and hence productivity, then we would expect to find a limited or absent relationship between CPP distance to countries that are *not* centers of biotechnology development and productivity. This idea motivates a falsification exercise, in which we re-estimate Equation 5.2, replacing $\text{CPPMismatchFrontier}^{\text{US}}_{k,\ell}$ with CPP mismatch with each country in the world; this generates a series of coefficient estimates $\hat{\beta}^{\ell}$, one for each country. That is, we estimate:

$$y_{k,\ell'} = \beta^{\ell} \cdot \text{CPP Mismatch}^{\ell}_{k,\ell'} + \chi_{\ell'} + \chi_k + \Omega'_{k\ell'}\Gamma + \varepsilon_{k,\ell'}$$

for all $\ell$. Figure A4 reports histograms of estimates of the $\hat{\beta}^{\ell}$, from specifications that do not include CPP mismatch the US as a control (Figure A4a) as well as from specifications that do (Figure A4b). In both cases, the coefficient on CPP mismatch with the US, marked with a dark green, dotted line, is the negative coefficient with the highest magnitude. Estimates of the effect of CPP distance to other countries are all smaller in magnitude and clustered around zero, especially conditional on CPP distance to the US.

Moreover, the $\hat{\beta}^{\ell}$ are significantly negatively correlated with country-level biotechnology development measured in the UPOV database. We estimate:

$$y_{\ell} = \xi \cdot \hat{\beta}^{\ell} + \varepsilon_{\ell}$$

where the dependent variable is either the number of varieties development in $\ell$ in the UPOV data, or an indicator that equals one of country $\ell$ enforces intellectual property protection for plant biotechnology. These are two proxies for the R&D intensity of country $\ell$. Table A8 reports estimates of $\xi$. The coefficient estimates are negative and significant, suggesting that CPP mismatch has more bite on global production for precisely the countries that are more active in R&D. These findings are consistent with our main estimates capturing the causal impact of technology's inappropriateness.

[46] These country-level characteristics include income, openness to trade, inequality, specialization in agriculture, agricultural productivity, and R&D intensity.

Table 4: CPP Mismatch Reduces Agricultural Output: Sub-national Estimates

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Dependent Variable is log Output | | | | | | | |
| | CPP Mismatch with the US | | | | CPP Mismatch with the Estimated Frontier | | | |
| CPP Mismatch (0-1) | -8.925*** | -10.20*** | -8.695*** | -9.355*** | -11.89*** | -10.10*** | -11.85*** | -10.37*** |
| | (2.386) | (3.327) | (1.752) | (2.096) | (1.937) | (2.475) | (1.538) | (2.247) |
| log(FAO-GAEZ-Predicted Output) | | 0.654*** | | | | 0.659*** | | |
| | | (0.138) | | | | (0.133) | | |
| *Included in LASSO Pool:* | | | | | | | | |
| Top CPP Fixed Effects | - | - | Yes | Yes | - | - | Yes | Yes |
| Ecological Features x Crop Fixed Effects | - | - | No | Yes | - | - | No | Yes |
| Crop x Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,436 | 696 | 1,437 | 1,093 | 1,370 | 696 | 1,371 | 1,036 |
| R-squared | 0.641 | 0.680 | | | 0.658 | 0.683 | | |

*Notes:* The unit of observation is a state-country pair. Columns 1-4 use CPP mismatch with the US and columns 5-8 use CPP mismatch with the estimated set of technological leader countries. Columns 1-2 and 5-6 report OLS estimates and columns 3-4 and 7-8 report post double LASSO estimates. State and crop-by-country fixed effects are included in all specifications, and included in the amelioration set in thet post-double LASSO specifications. The Top CPPs are defined as the top 200 CPPs defined by (i) the number of countries in whch they are present and (ii) the number of host crops that they infect. Since the two sets overlap, the total number is 335. The set of ecological features includes: temperature, precipitation, elevation, ruggedness, growng season days, soil acidity, soil clay content, soil silt content, soil coarse fragment volume, and soil water capacity. Standard errors are double-clustered by crop and state and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

### 5.5 Within-Country Estimates: Brazil and India

Finally, we exploit *state*-level information on CPP presence for both Brazil and India, along with the fact that both countries report detailed data on crop production at the state-level, to measure the effects of inappropriateness at a sub-national level. Our estimation framework is:

$$y_{k,s} = \beta \cdot \text{CPPMismatchFrontier}_{k,s} + \chi_s + \chi_{k.\ell(s)} + \Omega'_{k.s}\Gamma + \varepsilon_{k,s} \tag{5.3}$$

where now $s$ indexes states and $\ell(s) \in \{\text{Brazil}, \text{India}\}$. In all specifications, we include crop-by-country fixed effects ($\chi_{k.\ell(s)}$). By estimating the effect of inappropriateness on sub-national regions, we hold fixed all country-by-crop characteristics, including crop-specific R&D, trade, market size, demand, and pest composition. Thus, we estimate a qualitatively different parameter from the preceding analysis but also fully absorb potential unobservable features in the country-by-crop level analysis.

Our estimates of Equation 5.3 are displayed in Table 4, which follows the exact same structure as the baseline country-by-crop estimates in Table 3. Despite the inclusion of *country-by-crop* fixed effects, we find negative and significant estimates that are very similar in magnitude to our country-by-crop results. The coefficient estimates, if anything, increase when we account for local suitability, either controlling for state-by-crop level FAO GAEZ predicted output (columns 2 and 5), or using our more flexible post double LASSO approach (columns 3-4, 7-8). The findings are also very similar if we focus on either India or Brazil separately (Figure A5). Together, these estimates suggest that the (in)appropriateness of technology not only shapes productivity differences across country-crop pairs, but also shapes productivity differences across regions *within* countries for a given crop.

## 6. Case Studies: Inappropriateness and Technology Adoption

The empirical results of Sections 4 and 5 quantified the impact of CPP mismatch on technology diffusion and its consequences for production and specialization. In this section, we provide additional empirical evidence about the key intervening mechanism: that inappropriate technology is less likely to be adopted by farmers.[47] To do this, we home in on the geographically heterogeneous penetration of improved high-yielding varieties developed in the Green Revolution, and the relatively low usage of frontier agricultural technology in modern Africa.

### 6.1 High-Yield Varieties in the Green Revolution

The Green Revolution was a coordinated international effort, backed by philanthropic organizations like the Rockefeller Foundation, to develop high-yielding varieties (HYVs) of staple crops for countries with high risk of famine (Pingali, 2012). The engine at the heart of the Green Revolution was a set of international agricultural research centers (IARCs), including the International Rice Research Institute (IRRI) in the Philippines and the International Maize and Wheat Improvement Center (CIMMYT) in Mexico. These centers ultimately coalesced to form the Consultative Group for International Agricultural Research (CGIAR), an organization charged with coordinating international crop development for the poor world (Evenson and Gollin, 2003b).

Modern variety adoption and productivity growth during this period, however, still differed markedly across crops and countries (Evenson, 2005). One potentially important source of this heterogeneity, highlighted by scholars, is that varieties developed at the IARCs were inappropriate in places that are ecologically dissimilar from the countries in which the IARCs were located (Binswanger and Pingali, 1988; Lansing, 2009; Pingali, 2012). Lansing (2009) provides an in-depth case study of the detrimental impacts of the introduction of Green Revolution rice varieties and farming practices in Bali, where local practices had evolved to keep the local pest population at bay. The shift to Green Revolution technology precipitated widespread crop failures, driven by pest outbreaks.[48]

To investigate whether the inappropriateness of Green Revolution technology shaped its impacts, we first identify from Evenson and Gollin (2003b) the IARC and hence country in which breeding investment for each crop was centered (see Table A9). Using this information, we compute a measure of CPP mismatch with centers of Green Revolution breeding at the crop-by-country level:

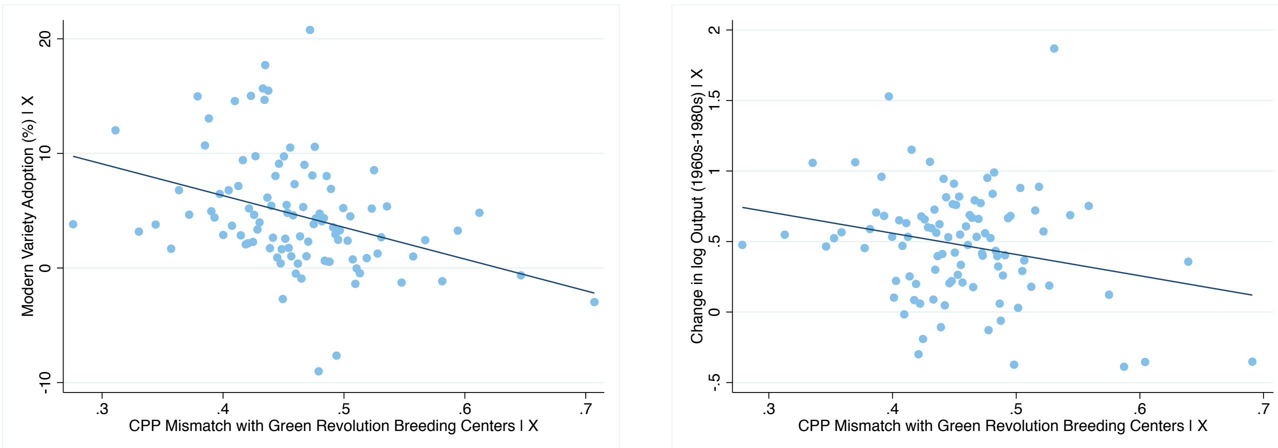
$$\text{CPPMismatchGR}_{k,\ell} = \sum_{\ell'} \text{CPP Mismatch}_{k,\ell,\ell'} \cdot \mathbb{I}\left\{\text{IARC for } k \text{ is in } \ell\right\} \tag{6.1}$$

[47] A formal articulation of this prediction is given in Corollary 1 in Appendix A.4.

[48] Reynolds and Borlaug (2006) document the significant challenges faced by CIMMYT to develop semi-dwarf wheat that would thrive outside of Mexico.

Figure 4: Inappropriateness and the Efficacy of the Green Revolution

(a) Modern Variety Adoption (b) $\Delta$ log Output (1960-1980)




*Notes*: This figure displays binned partial correlation plots, after absorbing country and crop-by-continent fixed effects, in which the independent variable is $\text{CPPMismatchGR}_{k,\ell}$ and the dependent variable is listed at the top of each sub-figure. In Figure 4a, the dependent variable is the share of production using modern varieties in 1980 ($p = 0.006$) and in Figure 4b, it is the change in log output between the 1960s and the 1980s ($p = 0.017$). Standard errors are clustered by country and continent-crop.

where $\mathbb{I}\{\text{IARC for } k \text{ is in } \ell\}$ is an indicator that equals one if Green Revolution breeding of crop $k$ was centered in country $\ell'$. For example, IRRI in the Philippines was the main IARC for rice, so in all countries CPP distance for rice is computed as CPP distance to the Philippines.

We first study the relationship between $\text{CPPMismatchGR}_{k,\ell}$ and modern variety adoption at the crop-by-country as reported by Evenson and Gollin (2003a,b). We regress the percent of area devoted to high-yield varieties in 1980-85, a representative cross-section after the bulk of Green Revolution research was instigated, on $\text{CPPMismatchGR}_{k,\ell}$ and absorbed effects at the location and crop-by-continent ($k \times c(\ell)$) level:

$$\text{HYVAdoption}_{k,\ell,1980} = \beta \cdot \text{CPPMismatchGR}_{k,\ell} + \chi_\ell + \chi_{k,c(\ell)} + \varepsilon_{k,\ell} \tag{6.2}$$

Our sample is the 8 crops in Table A9 intersected with the 85 low-income countries in the Evenson and Gollin (2003a,b) data.

CPP mismatch with centers of Green Revolution breeding substantially reduced the adoption of modern seed varieties. The main finding is summarized in Figure 4a, which shows a binned partial correlation plot of $\beta$ estimated from Equation 6.2. Our estimate of $\hat{\beta} = -26.62$ (9.15) implies that the 75th percentile crop-country pair had 5 percentage points lower HYV penetration than the 25th percentile in 1980, relative to a mean HYV penetration value of 5%. If we restrict attention to corn, wheat, and rice, the three most prominent Green Revolution crops, our coefficient estimate jumps to

$\hat{\beta} = -96.20$ (27.17), implying a 18 percentage point difference between the 75th and 25th percentiles relative to a mean of 14% (see Table A10).[49]

We next directly study the impact of this heterogeneous adoption on production and specialization by adapting our empirical framework from Section 5. In particular, we study how CPP mismatch with Green Revolution centers affected output *growth* from the 1960s to the 1980s, the period when the majority of Green Revolution research took off. We estimate the following regression model:

$$y_{k,\ell,1980s} - y_{k,\ell,1960s} = \beta \cdot \text{CPPMismatchGR}_{k,\ell} + \tau \cdot y_{k,\ell,1960s} + \chi_\ell + \chi_{k,c(\ell)} + \varepsilon_{k,\ell} \tag{6.3}$$

where the dependent variable is the *change* in (log of) crop-level output between the 1960s and the 1980s, and the sample includes all crop-country pairs from the HYV adoption model. This estimating equation differences out the direct effects of time invariant ecology and local suitability, identifying how changes in output respond to changes in the geography of innovation (and hence inappropriateness) relative to the relevant set of innovating countries.

Our finding, summarized as a binned partial correlation plot in Figure 4b, is that production shifts away, in relative terms, from crop-location pairs more ecologically mismatched with Green Revolution hubs. Our coefficient estimate $\hat{\beta} = -2.64$ is about 1/3 of our previously estimated point estimate for the effects of modern inappropriateness relative to the technological frontier.[50] Table A11 documents that the relationship between CPP distance to Green Revolution breeding centers and changes in production is restricted to the period 1960-1980, the height of the Green Revolution (columns 1-3); the effect is apparent in Asia, Africa, and South America, but not in Europe, which was not an intended recipient of Green Revolution technology (columns 4-7). These findings are consistent with a causal interpretation of the main result.

Taken together, our findings illustrate how geographical inappropriateness shaped impact of the Green Revolution and, more broadly, how changes in the centers of innovation can shift the relationship between ecological conditions and productivity. The focus of the Green Revolution on developing a relatively small set of HYVs and distributing them widely may have undermined its global reach, since new varieties were less productive and less likely to be adopted in the first place in environments that were different from HYV breeding centers.

### 6.2 Technology Adoption in Sub-Saharan Africa

We next study how inappropriateness affects production on smallholder farms in sub-Saharan Africa, which have received substantial attention for the low penetration of agricultural technology in spite of ostensible benefits (see, e.g., Suri, 2011; Duflo et al., 2011). Our specific question is the extent to which the inappropriateness of frontier technology explains low use of improved inputs.

[49] In a falsification exercise, we estimate the relationship between HYV adoption and CPP distance to all other countries, and we compile these placebo coefficients. Our main estimate is in the far left tail of the coefficient distribution ($p = 0.013$), indicating that our findings are truly driven by features of IARC ecology and not spurious correlation.

[50] Table A10 reports summarizes the estimates for both of our regression models.

To measure the use of improved technologies, we combine data from the latest round of all Living Standard Measurement Survey (LSMS) Integrated Surveys of Agriculture (ISA). These are detailed surveys on all facets of agricultural production, including technology use, collected by the World Bank in collaboration with the statistical agencies of eight countries: Burkina Faso, Ethiopia, Malawi, Mali, Niger, Nigeria, Tanzania, and Uganda. Data are collected at the field and farm level, and the LSMS-ISA also provides the coordinates of the approximate location of each farm.[51] Our key dependent variable of interest is farm-by-crop information on the use of improved seeds (i.e., not locally bred or "traditional" varieties). We construct an indicator variable for each crop grown in each farm if improved seed varieties are used. In total, we have data on approximately 120,000 crop-farm pairs across all eight countries.

Our main estimating equation is:

$$\text{ImprovedSeed}_{k,z} = \beta \cdot \text{CPPMismatchFrontier}_{k,\ell(z)} + \chi_{\ell(z)} + \chi_k + \varepsilon_{k,z} \tag{6.4}$$

where $k$ continues to index crops and $z$ indexes farms in the LSMS-ISA data. The dependent variable is an indicator that equals one if farmer $z$ uses an improved seed variety for crop $k$. $\chi_k$ denote crop fixed effects and $\chi_{\ell(z)}$ denote country fixed effects, included in all specifications. If the inappropriateness of technology reduces technology adoption, we would expect that $\beta < 0$; however, it is possible that the smallholder farmers in the sample are not likely to use improved technology regardless of its appropriateness, and the context specificity of frontier innovation is not an important barrier to productivity enhancements in this setting.

Our findings are reported in Table 5, where CPP mismatch is measured either as CPP mismatch with the US (Panel A) or CPP mismatch with the measured set of crop-specific frontier countries. We estimate a negative and significant relationship between adoption and CPP mismatch. The estimates of column 1 imply that improved seed use by the median farmer in our sample would be 14% more prevalent absent inappropriateness, relative to an in-sample mean of 17.9%. The estimates are similar after including state fixed effects (column 2) or a quadratic polynomial in farm latitude and longitude (column 3) in order to control flexibly for the effect of geography. Our findings are also similar when the regression is weighted by farm size (column 4) or using our two alternative constructions of CPP mismatch (columns 5-6; these use the "broad" CPP presence definition and Jaccard functional form)

These estimates indicate that inappropriateness contributes toward low improved input use on some of the world's least productive small farms. Through the lens of our model, in which endogenous innovation responds to demand for inputs, they further suggest a reason why research and marketing investment from global biotechnology firms has not materialized in sub-Saharan Africa (Access to Seeds Foundation, 2019), despite the ostensibly large market opportunity.

[51]To preserve farmer anonymity, the LSMS-ISA provides the latitude and longitude of each survey cluster rather than unique coordinates for each household. To keep a consistent sample across specifications, we restrict our analysis to households in which the cluster coordinates were included in the data set.

Table 5: CPP Mismach Inhibits Biotechnology Adoption in Africa

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Dependent Variable is Improved Seed Use (=1) | | | | | |
| | *Panel A: CPP Mismatch with the US* | | | | | |
| CPP Mismatch (0-1) | -0.220*** | -0.186*** | -0.185*** | -0.147*** | -0.205*** | -0.314*** |
| | (0.0635) | (0.0610) | (0.0614) | (0.0511) | (0.0689) | (0.0870) |
| Observations | 115,397 | 115,393 | 115,393 | 104,623 | 115,393 | 115,393 |
| R-squared | 0.213 | 0.246 | 0.247 | 0.235 | 0.247 | 0.247 |
| | *Panel B: CPP Mismach with the Estimated Frontier Set* | | | | | |
| CPP Mismatch (0-1) | -0.321*** | -0.242*** | -0.237*** | -0.157*** | -0.227*** | -0.237*** |
| | (0.0793) | (0.0805) | (0.0812) | (0.0563) | (0.0793) | (0.0812) |
| Observations | 114,605 | 114,601 | 114,601 | 103,968 | 114,601 | 114,601 |
| R-squared | 0.213 | 0.246 | 0.247 | 0.235 | 0.246 | 0.246 |
| Quadratic Polynomial in Lat and Lon | | | ✓ | ✓ | ✓ | ✓ |
| log Area-Weighted Estimates | | | | ✓ | | |
| Broad CPP Presence Classification | | | | | ✓ | |
| Jaccard (1900, 1901) Mismatch Metric | | | | | | ✓ |
| Crop Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | - | - | - | - | - |
| State Fixed Effects | No | Yes | Yes | Yes | Yes | Yes |

*Notes:* The unit of observation is a plot. In Panel A, CPP mismatch with the frontier is estimated as CPP mismatch with the US and in Panel B it is estimated using the frontier set selected from the UPOV data. The controls included in each specification, as well as the mismatch metric when the baseline measure is not used, are noted at the bottom of each column. Standard errors are clustered by crop-country and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

# 7. Inappropriate Technology and Productivity: Present and Future

In this section, we use empirical estimates from Section 5 in combination with the model to study how the inappropriateness of technology and existing ecological bias of the global innovation system affects global productivity. We first explicitly describe the mapping from our empirical results to our model interpretation and counterfactuals (Section 7.1). We then the level and distributional effects of "removing inappropriateness" in the observed equilibrium (Section 7.2). We finally use our model to study the productivity effects of three potential future scenarios: targeted research in a Second Green Revolution (Section 7.3), the realignment of agricultural research toward emerging markets (Section 7.4), and the global movement of crop pests and pathogens due to climate change (Section 7.5).

## 7.1 Methods

### 7.1.1 From Theory to Data

Our empirical findings about technology transfer in Section 4 and production distortions in Section 5 suggest that the observed world equilibrium is well-approximated with a structure of a few "leaders" driving the frontier of agricultural technology. In this subsection, we describe a simplification of our full model from Section 2 which embodies this logic, maps transparently to the empirical findings, and allows us to formally define counterfactual scenarios of interest.

Concretely, we specialize the model by assuming that for each crop $k$ there is a "Frontier technology producer" $F(k) \in \{1, \ldots, L\}$. In the Frontier producer of each crop $k$, general research is inelastically supplied at level $\bar{A}(k) > 0$, own-CPP research at level $\bar{B} > 0$, and foreign-CPP research at level $\bar{B}e^{-\hat{\tau}}$ for some $\hat{\tau} > 0$.[52] These assumptions encode a fixed knowledge gap in productivity units for each crop, to match our empirical identification strategy. They abstract from the endogeneity of the magnitude of knowledge gaps in response to incentives, a topic about which we have little information in the data. We finally close the model in general equilibrium by assuming each crop price $p(k)$ lies on the isoleastic demand curve

$$\frac{p(k)}{\bar{p}(k)} = \left(\frac{Y(k)}{\bar{Y}(k)}\right)^{-\varepsilon} \tag{7.1}$$

where $(\bar{p}(k), \bar{Y}(k))_{k=1}^{K}$ are constants, $Y(k)$ is total production of crop $k$, and $\varepsilon > 0$ is an elasticity of demand for each crop relative to a numeraire (e.g., a good representing the rest of the economy). This model recognizes that international prices provide a natural hedge against lower physical productivity, but abstracts from specific patterns of demand substitution across crops.

We now describe the key model predictions about specialization and productivity, introduced in Proposition 2, in the context of this case of the model. Let $\delta(k, \ell, F(k))$ denote CPP dissimilarity with the crop-specific frontier. Production of crop $k$ in country $\ell$ is given by

$$\log Y(k, \ell) = -\eta\gamma\delta(k, \ell, F(k)) + \eta(\log p(k) + \log \omega(k, \ell) + \alpha\bar{A}(k) + (1-\alpha)\bar{B}) - (\eta - 1)\log \Xi(\ell) \tag{7.2}$$

where $\gamma := (1-\alpha)\hat{\tau} > 0$ is the sensitivity of log crop-specific productivity to CPP dissimilarity in the model and $\Xi(\ell)$ is the productivity index:

$$\log \Xi(\ell) = \alpha \log \bar{A}(k) + (1-\alpha)\log \bar{B} + \frac{1}{\eta}\log\left(\sum_{k=1}^{K} p(k)^{\eta}\omega(k,\ell)^{\eta}e^{-\eta\gamma\delta(k,\ell,F(k))}\right) \tag{7.3}$$

Comparing Equation 7.2 with the regression model Equation 5.2 reveals that our empirical estimate of $\beta$, the sensitivity of log output to CPP dissimilarity, identifies the product of the productivity effect $\gamma$ and the elasticity of supply $\eta$. Equation 7.3 shows how, conditional on separately identifying $\gamma$ (the direct productivity effect) and $\eta$ (the elasticity of supply), we can translate our estimates into total country-level revenue productivity.

In the next section, will elaborate on exactly how we will calibrate the model to incorporate each of these forces. We first precisely define how we will conduct counterfactual analysis in the context of the present model. We describe a counterfactual scenario that "removes inappropriateness" as one in which non-local-CPP research is subsidized to reach level $\bar{B} > \bar{B}\exp(-\hat{\tau})$ in all frontier countries.

[52] More formally, in the frontier countries, we set $B_0 = \bar{B}^{-1}$ and take a limit of $\phi \to \infty$ and $\tau \to \infty$ such that $\frac{\tau(\bar{B})}{1+\phi} \to \hat{\tau} > 0$. In other countries, we set $B_0 \to \infty$ so no research is performed.

Table 6: Model Parameters and Data for Estimation

| Parameter | Estimate | Specification/Source | Definition |
|---|---|---|---|
| $\beta$ | -7.14 | Equation 5.2 | Reduced form effect of CPPDistFrontier on output |
| $\eta$ | 2.46 | Costinot et al. (2016) | Elasticity of supply to productivity |
| $\gamma$ | 2.90 | $-\beta/\eta$ | Sensitivity of log productivity to CPPDistFrontier |
| $\varepsilon$ | 0.35 | Muhammad et al. (2011) | Price elasticity of global food demand |
| $\pi(k,\ell)$ | — | FAOSTAT Database | Planted area for each crop in each country |
| $\Xi(\ell)$ | — | Fuglie (2012, 2015) | Baseline total revenue productivity by country |

This intervention removes the knowledge gap between frontier and non-frontier CPP research by replicating the missing knowledge spillover, and it undoes the depressive effect of CPP differences on technology diffusion. While we make no claim that such an intervention is "optimal" in the underlying model under a welfare criterion, it provides one reasonable and interpretable benchmark for the total productivity effect of the "inappropriate technology bias." This counterfactual scenario will our focus in Section 7.2, and the blueprint for defining all subsequent counterfactual experiments.

Letting hats denote quantities under the "removal of inappropriateness" scenario, it is straightforward to show that changes in production are given by

$$\log \hat{Y}(k,\ell) - \log Y(k,\ell) = \eta\gamma\delta(k,\ell,F(k)) + \eta(\log \hat{p}(k) - \log p(k)) - (\eta - 1)\left(\log \hat{\Xi}(\ell) - \log \Xi(\ell)\right) \quad (7.4)$$

and changes in revenue productivity by

$$\log \hat{\Xi}(\ell) - \log \Xi(\ell) = \frac{1}{\eta}\log\left(\sum_{k=1}^{K} \hat{p}(k)^{\eta}\omega(k,\ell)^{\eta}\right) - \frac{1}{\eta}\log\left(\sum_{k=1}^{K} p(k)^{\eta}\omega(k,\ell)^{\eta}e^{-\eta\gamma\delta(k,\ell,F(k))}\right) \quad (7.5)$$

Changes in productivity arise from a partial-equilibrium effect of removing the depressive effect of inappropriateness and a general-equilibrium effect of price adjustment.

### 7.1.2 Calibration

As alluded to above, measuring the productivity effect of inappropriateness involves additional information about the elasticity of supply to productivity changes. Our strategy is to obtain an external estimate of the supply elasticity ($\eta = 2.46$) from Costinot et al. (2016), who study productivity changes and re-allocation in global agricultural production using the same Fréchet discrete choice model.[53,54] Combining this estimate with our baseline estimate of $\beta = -7.14$ (Table 3, column 5) yields an estimate of $\gamma = 2.90$, in units of percent productivity loss per basis point of CPP distance.

Conditional on $\eta$, the crop-by-location productivity $\Theta(k,\ell)$ is identified up to scale from data on

[53] Sotelo (2020), studying Peruvian agriculture, finds a comparable estimate of $\eta = 2.06$.

[54] These authors' estimate, in a nutshell, is the plot-level heterogeneity required to explain the relationship between agronomically measured productivity (from the FAO-GAEZ model) and observed planting patterns at the plot level (about 50-square-kilometer-size) in the modern world.

Table 7: Causal Effects of Inappropriateness

| Statistic | Unit | Scenario Flexible Prices | Fixed Prices |
|---|---|---|---|
| Reduction in Productivity | Percent | 42.2 | 56.5 |
| | | (4.0) | (4.7) |
| Increase in IQR (75-25) | Percent | 15.1 | 19.7 |
| | | (0.4) | (0.7) |

*Notes*: Calculations compare the observed equilibrium with inappropriate technology to the counterfactual equilibrium without inappropriate technology, as described in the main text. Standard errors, in parentheses, are calculated via the Delta Method, using the numerical gradient of statistics to the estimated parameter $\beta$. Productivity losses are area-weighted means across countries.

planted area by crop, $\pi(k, \ell)$.[55] Mirroring our analysis in Section 5, we measure these areas using the crop-by-country planting data from the FAOSTAT database, averaged from 2000-2016. To maximize clarity when reporting country-level results, and to limit the effect of outlier observations, we ignore countries in the bottom decile of total agricultural area (the largest such country is Mauritius).

We use estimates of total agricultural revenue from Fuglie (2012, 2015), again averaged from 2000 to 2016, to calibrate all countries' initial revenue productivity and hence pin down the scale of local innate productivity and prices. In our results, unless otherwise stated, we define "productivity" as productivity per acre. Finally, to calibrate the crop-level demand curves, we use the average value estimated by the US Department of Agriculture for the (compensated) own-price elasticity of global food consumption (Muhammad et al., 2011). This yields $\varepsilon = 0.35$.[56] All necessary model parameters are listed and summarized in Table 6.

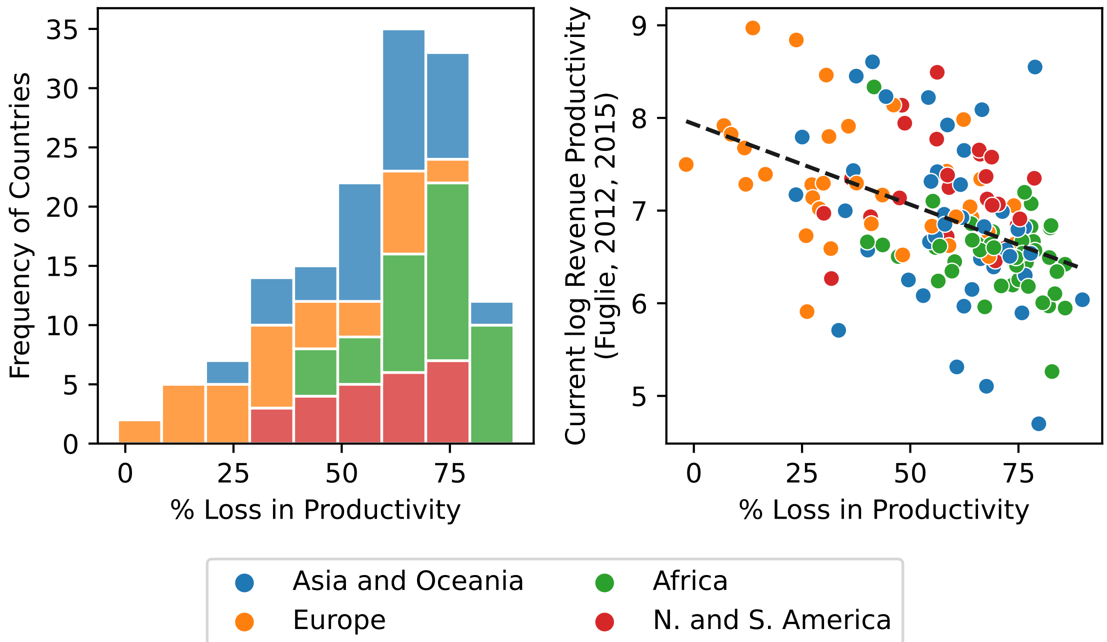
### 7.2 The Productivity Effects of Inappropriateness

We first study the counterfactual scenario of removing inappropriateness. In Table 7, we summarize our main findings about productivity and productivity disparities in the *observed equilibrium relative to the counterfactual equilibrium*. We report average productivity changes (area weighted averages) for the entire world, and the percent change in the 75-25 percentile gap (inter-quartile range) of log productivity. In our baseline model with price adjustment, inappropriateness reduces global productivity by 42.2%. Inappropriateness increases the IQR of the log productivity distribution by 15.1%; in other words, inappropriateness "explains" this percentage of global disparities. We also report results under an alternative model with fixed prices ($\varepsilon = 0$) to gauge the importance of the global price hedge. As expected, the effects under rigid prices are larger in the aggregate and for reduction of disparities.

[55]The model suggests that an equivalent method is to use production in value terms. We favor using areas because it avoids the need for data on local prices.

[56]Specifically, we use the average of the "low," "middle," and "high" income estimates in Appendix Table 3 of that publication.

Figure 5: Causal Effects of Inappropriateness: Heterogeneity by Location




*Notes*: The left graph is a histogram of productivity losses from inappropriateness. The right graph is a scatterplot of productivity losses against observed productivity. The dashed line is a best-fit linear regression across countries. In each plot, colors indicate continents.

We next more closely explore the distributional implications of our findings. The left panel of Figure 5 shows the distribution of productivity losses by country as a histogram, focusing on the full-model, flexible-price calculation. The largest losses from inappropriateness are concentrated in Africa and Asia, while the smallest are in Europe. The right panel documents a negative and significant relationship between our estimated productivity losses and present-day revenue productivity (coef. $= -0.017$, $t = -6.2$). Thus, inappropriateness has the largest negative effects on productivity in precisely the countries that are least productive today.[57]

These results, taken together, put into sharp relief the inequality created by the interaction of ecological heterogeneity and the global innovation system. Neglected agricultural ecosystems, like neglected tropical human diseases (Hotez et al., 2007), are concentrated in specific and predominantly poor parts of the world. These regions are unproductive today, and kept unproductive due to an absence of appropriate technology or incentives to develop it. Our framework suggests that the main short-run remedies are policies that seed the ground for local biotechnological research.

**Sensitivity Analysis.** Our empirical analysis is focused on accurately estimating $\beta$, the effect of CPP mismatch on output. As discussed above, in order to estimate the aggregate effects of inappropriateness we also rely on two additional parameters that we obtain from existing literature, the elasticity

[57]Some, but not all, of this effect is spanned by the cross-continent variation highlighted above. Replicating the same regression model with continent fixed effects gives a coefficient of -0.014 and $t$-statistic of -3.8.

of supply to productivity ($\eta$) and the price elasticity of food demand ($\varepsilon$). To explore sensitivity of our findings, we identify maximum and minimum plausible estimates of each parameter from the literature. For the maximum and minimum plausible values for $\varepsilon$, we use the maximum and minimum price elasticities reported in Muhammad et al. (2011). For the minimum plausible value for $\eta$, we use $\eta = 2$ which is slightly lower than the estimate of $\eta = 2.06$ in Sotelo (2020), to our knowledge the lowest estimate of the relevant parameter in existing literature. For the maximum plausible value, we add the difference between the Sotelo (2020) estimate and our baseline estimate of $\eta$. Our results are reported in Figure A9, which recreates the histogram of losses and relationship between current productivity and losses as shown for our baseline in Figure 5. Our findings of large average losses, between 30% to 60% of counterfactual agricultural productivity, and significantly greater losses in observed unproductive locations are robust across parameter choices. As expected, reducing price impacts (increasing $\varepsilon$) dampens the effects of inappropriateness, while decreasing the extent of unobserved heterogeneity (decreasing $\eta$) amplifies them.

**Inappropriateness Due to Other Ecological Differences.** Our main results focus on CPP mismatch as a key shifter of technology diffusion and inappropriateness. However, as highlighted in Section 3.4, CPP mismatch is not the *only* determinant of inappropriateness; other features of ecological and geographic mismatch with the frontier could contribute to the inappropriateness of modern technology and affect the aggregate effect of inappropriateness on global productivity. In Section B.2, we describe our measurement of ecological mismatch in non-CPP related features, including temperature, precipitation, topography, and soil characteristics. We then estimate the effect of non-CPP ecological mismatch with the frontier on output, and study the counterfactual scenario of removing inappropriateness in the form of this broader set of geographic and ecological features, in addition to CPP mismatch. Figure A10 shows the equivalent of Figure 5, visualizing the cross-country distribution of losses due to inappropriateness, under this scenario. At the aggregate level, incorporating these additional dimensions of potential inappropriateness increases our estimate of the losses due to inappropriateness total productivity to 52%, and increases the effect on disparities in productivity to 16%. Comparing these estimates to our baseline reported in Table 7, we find that CPP mismatch has roughly four times the effect on global output as the combination of all non-CPP characteristics. This finding further justifies our focus on CPP mismatch for the main analysis.

### 7.3 Mapping a Second Green Revolution

Having studied how the present distribution of biotechnology research shapes global productivity, we now use our model to study the effects of counterfactually shifting that distribution. Our first exercise, in the spirit of the historical Green Revolution, is to study how to target a modern "Second Green Revolution" that is as appropriate, and as productivity enhancing, for the world as possible.

Concretely, for each of the eight major crops that were the focus of the historical Green Revolution, we calculate the counterfactual productivity benefit of moving the "Frontier" to each possible country

Table 8: Inappropriateness-Minimizing Centers for Modern Agricultural Innovation

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|
| | Sites Chosen to Minimize Global Inappropriateness | | | | Sites Chosen to Minimize Inappropriateness in Countries with Below Median Productivity | | | |
| Crop | Best Site | % Change in Productivity | Second Best Site | % Change in Productivity | Best Site | % Change in Productivity | Second Best Site | % Change in Productivity |
| Wheat | China | 3.29 | India | 1.87 | India | 10.42 | Pakistan | 6.97 |
| Maize | China | 8.50 | USA | 6.16 | Nigeria | 9.26 | Tanzania | 7.46 |
| Sorghum | India | 0.83 | Nigeria | 0.76 | Nigeria | 3.10 | India | 2.71 |
| Millet | Nigeria | 0.90 | India | 0.68 | Nigeria | 2.97 | Zimbabwe | 1.76 |
| Beans | India | 1.30 | Brazil | 1.13 | India | 3.25 | Tanzania | 1.41 |
| Potatoes | China | 0.97 | India | 0.48 | India | 0.94 | Russia | 0.52 |
| Cassava | Nigeria | 0.41 | Ghana | 0.31 | Nigeria | 1.60 | DRC | 1.33 |
| Rice | China | 7.55 | India | 6.53 | India | 13.32 | Thailand | 8.65 |

*Notes:* Column 1 reports the crops included in our analysis of the Green Revolution. Columns 2-5 report the results of our analysis to select the two countries where breeding investment would have the largest positive effect on global output for each crop. Columns 6-9 report the results of our analysis to select the two countries where breeding investment would have the largest positive effect on output in countries with below median productivity for each crop. All estimates rely on the full model with non-linear adjustments and price responses.

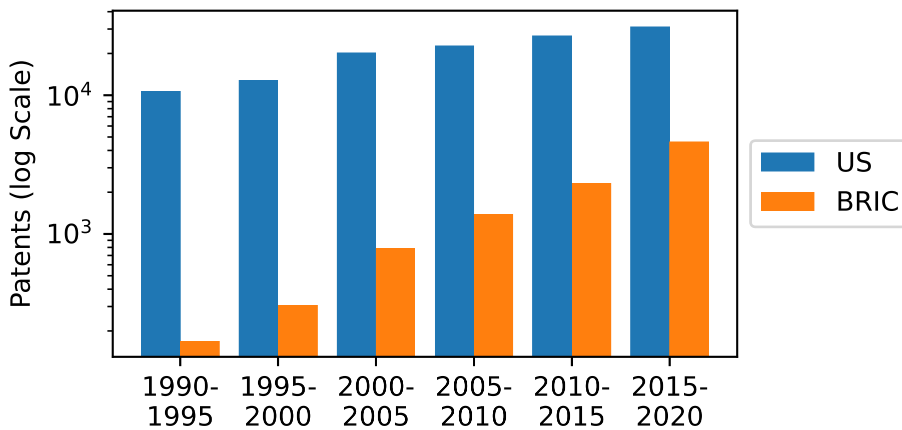
$\ell' \in \{1, \ldots, L\}$. As in our previous exercise, we consider this as a pure adjustment to inappropriateness without shifting the maximum productivity of frontier research or the size of knowledge gaps, which are controlled by $(\bar{A}, \bar{B}, \hat{\tau})$. We identify the new Frontier choices that would have the largest effect on global productivity, as well as on productivity in initially below-median-productivity countries.

Table 8 displays the results for each studied crop. Columns 2 and 4 report the two countries where breeding research would increase global output by the most, and columns 3 and 5 show the corresponding quantitative effects on global agricultural productivity in log points times 100. Columns 6-10 report analogous results if we instead calculate productivity gains only for countries with below median productivity in the contemporary cross-section.

This analysis, while necessarily speculative, yields several interesting conclusions. First, the set of countries that increases total versus low-productivity countries' output is similar. This is consistent with our findings in Section 7.2 that reducing ecological mismatch would both increase global output and reduce production disparities. Second, the findings are consistent with the hypothesis that a lack of breeding in Africa, including during the Green Revolution, holds back global productivity growth (Pingali, 2012). Nigeria, Ghana, Zimbabwe, Tanzania, and the Democratic Republic of Congo all emerge as countries where breeding research could have large, positive effects.

Finally, the prominence of China on the lists highlights the role that geopolitics might have played and continue to play in shaping where research takes place. Political connections may limit where governmental or philanthropic organizations can invest in technology development, constraining the potential of such investments to develop globally appropriate technology. The same pattern, however,

Figure 6: Growth in Agricultural Patented Technologies, BRIC vs. the United States




*Notes*: Total number of patented agricultural technologies (i.e., in CPC class A01) in each five year period, comparing patents with assignees in the US to patents with assignees in Brazil, Russia, India, or China, from one of the five major patent offices (USPTO, WIPO, EPO, JPO, KIPO). Bars are the number of patents issued in the five year bin noted on the horizontal axis.

suggests that there are potentially large opportunities for countries like China, India, and Russia—growing players in global R&D—to market their technology around the world, especially in countries where appropriate technology is lacking today.

### 7.4 New Biotechnological Leaders

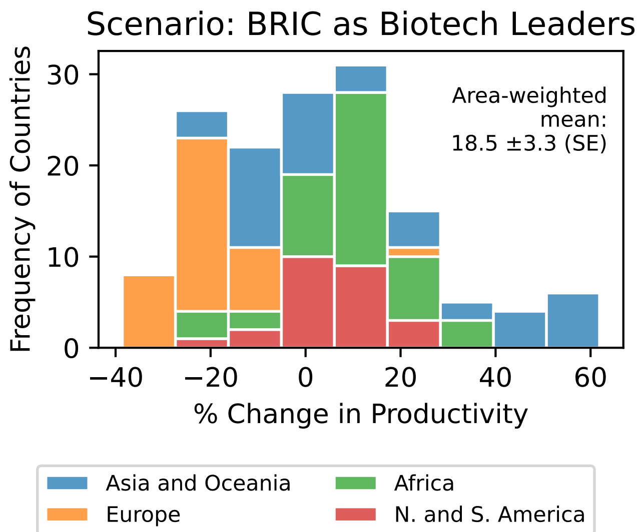
In the past several decades, the United States and Western Europe have been at the center of global biotechnology development. However, there is reason to believe that the landscape of biotechnology research could change in the coming years—and some evidence that this process has already begun. Figure 6 displays the number of patented agricultural technologies over time, relative to the period 1990-1995, comparing technologies developed in the United States to the trend for technologies developed in "BRIC" countries (Brazil, Russia, India, China). While throughout the period the *level* of innovation in the US is higher, agricultural innovation has grown substantially in the BRICs.

What might the impact in this shift in the center of global research be on global productivity? The findings of our Second Green Revolution exercise (Table 8) hinted that shift in international focus may be broadly beneficial for boosting global productivity and reducing disparities. Moreover, several anecdotes suggest that BRIC-nation policymakers have recognized the associated business—and soft power—opportunities from investment in agricultural R&D.[58]

To operationalize a "rise of BRIC" scenario in our model, we first calculate the CPP mismatch of

[58] As one example, the Brazilian Agricultural Research Corporation (EMBRAPA), a state-owned agricultural research organization, has a long-standing cooperation with several African countries based on the premise of their ecological similarity. For example, see here: https://www.embrapa.br/en/cooperacao-tecnica/m-boss. The description of the collaboration on the EMBRAPA website argues that the "exchange of knowledge and technologies is facilitated due to similarities in their cultures, climate, ecosystems, and agricultural practices."

Figure 7: Rise of BRIC: Global Productivity Changes




*Notes*: This graph reports a histogram of productivity changes in the counterfactual scenario where we simulate the rise of Brazil, Russia, India, and China (BRIC) as biotechnological leaders.

every country-crop pair with the BRIC research frontier as:

$$\text{CPPMismatchFrontier}^{\text{BRIC}}_{k,\ell} = \sum_{\ell' \in \text{BRIC}} \frac{\pi(\ell', k)}{\sum_{\ell'' \in \text{BRIC}} \pi(\ell'', k)} \times \text{CPPMismatch}_{k,\ell,\ell'} \tag{7.6}$$

In words, we estimate the inappropriateness of BRIC ecology for each crop, weighting each BRIC country by its share of total area devoted to that crop within the BRIC countries.[59] We then, analogously to the previous counterfactual experiments, consider the effects of moving the frontier such that $\delta(k, \ell, F(k)) = \text{CPPMismatchFrontier}^{\text{BRIC}}_{k,\ell}$.

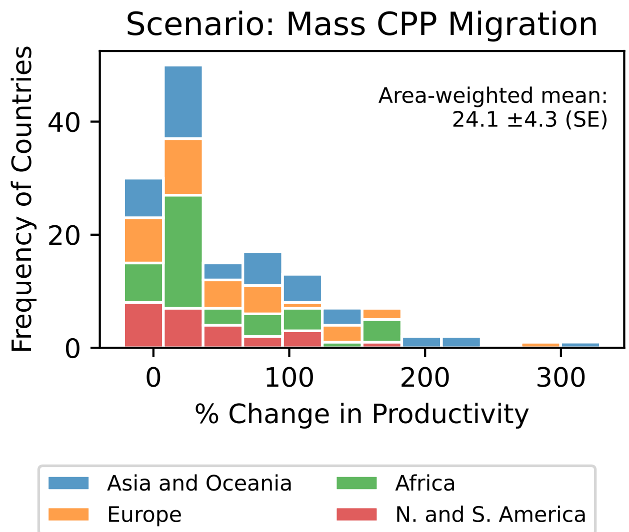
Figure 7 summarizes our findings in a continent-coded histogram of the implied revenue productivity changes. The average effect is a 18.5% productivity boost, speaking to the fact that the BRIC countries span more ecological diversity than the existing set of technological leaders. Africa stands particularly to gain, on average, from this realignment, even though none of the BRIC countries are in Africa itself. However, there are also clear losers, including several countries in Europe and Asia, which benefit from their ecological similarity to the current set of technological leaders. From the perspective of the developing world, a shift of innovation investment to the BRIC nations may be a partial, if incomplete, substitute for encouraging purely local technological development.

### 7.5 Ecological Differences Under CPP Mass Migration

So far, we have treated ecology as immutable and allowed innovation to move around the world. But climate change has accelerated changes in ecological systems themselves, and will continue to do so

[59] For crops that are not cultivated in any BRIC country, we use the estimated leader countries from the main analysis.

Figure 8: Climate-Induced CPP Migration: Global Productivity Changes




*Notes*: This graph reports a histogram of productivity changes in the counterfactual scenario where we simulate the future migration of CPPs due to climate change.

over the coming decades (Parmesan and Yohe, 2003). In the context of crop pests and pathogens, increases in temperature lead to a systematic, poleward movement (Bebber et al., 2013). While poleward CPP movement to date has been limited (Bebber et al., 2013), temperature change over the past fifty years is also much more limited than projected temperature change over the coming decades.[60] This could change the relevant "geography of innovation" by shifting the relevant set of CPP threats in each country, even if the identity of innovating countries remains fixed.

The impact of climate change on the appropriateness of frontier technology across crops and countries is also not clear *ex ante*. If CPP range shifts increase the CPP similarity between a given country and R&D intensive regions, then it might be able to more effectively make use of technology developed in the new equilibrium. However, CPP movement could also reduce the the CPP overlap across countries if, for example, the US inherits several unique CPPs from Central America (or Europe from North Africa), reducing their similarity to other large parts of the world. To capture this channel, we extrapolate the estimates in Bebber et al. (2013) of poleward CPP movement to date into the future, using projected changes in global temperature due to climate change between the present and 2100.[61] We then use these data to construct CPPDistFrontier$(k, \ell)^{\mathrm{CC}}$ based on ecological dissimilarity to the

[60] In the data, CPPs have moved poleward over the past 50 years by about 135 kilometers (Bebber et al., 2013). While global temperatures have increased by about 1°C over the past 50 years, in a "worst case" future scenario, temperature is projected to increase by 4.3°C by 2100. This projection corresponds to Representative Concentration Pathway (RCP) 8.5, a consensus worse-case scenario.

[61] The consensus worst case scenario implies a 4.3°C increase in temperature by 2100, and hence a 700km poleward movement of CPPs on average (or approximately the distance from Tunis to Rome). We simulate poleward range spread of each pest by identifying all countries that intersect a 700km translation of all countries that presently contain the CPP, and appending these matches to the observed presence data to construct a dataset of predicted CPP presence in 2100. Finally, we include manual corrections for countries with non-contiguous territory.

modern set of frontier innovators, and re-calculate productivity as in the previous counterfactuals.

Figure 8 shows that we find an overall positive effect, which is relatively evenly spread across space. Our analysis therefore highlights that increasing *ecological similarity* may provide a partially offsetting force to the directly negative effects of ecological change, insofar as it coordinates the global research system around a more common set of productivity threats. This dynamic in agricultural innovation, and in climate-induced innovation more broadly, is an important topic for further research.

## 8. Conclusion

We investigate a long-standing hypothesis that frontier technologies' endogenous *appropriateness* for the high-income countries that develop them shapes global patterns of technology diffusion and productivity. Our empirical focus is global agriculture, and we develop a new measure of the potential inappropriateness of crop-specific agricultural biotechnology based on the dissimilarity in crop pest and pathogen (CPP) environments across locations. We first show technology development is concentrated in a small set of countries and focused on local pest and pathogen threats. We next show that environmental mismatch is a substantial barrier to the international diffusion of crop-specific biotechnology, and that countries move production away from crops for which their CPP mismatch with the research frontier is higher. Technological progress in the frontier, far from diffusing broadly and evenly around the world, underlies global inequality.

Combining our estimates with a model of global agricultural production, we estimate that inappropriateness as proxied by CPP mismatch reduces global agricultural productivity by 40-55%, and increases global disparities in agricultural productivity by 10-15%. Substantial differences in pest and pathogen threats around the world, and innovators' neglect of ecosystem threats in low-income areas, sustains large disparities in access to appropriate technology and, as a result, in productivity. Moreover, changes in the geography of innovation can have large effects on the distribution of appropriate technology, and hence productivity, around the world. We show that the global impact of the Green Revolution was shaped by ecological similarities differences with the key breeding centers, and argue that in the future, changes in the center of global biotechnology development and in ecology due to global warming could shift features of the technological frontier and hence appropriateness of technology around the world. More exploration of these trends, which will define agriculture and technology in the coming century, is an important area for future research.

## References

Access to Seeds Foundation (2019). Access to Seeds Index: 2019 Synthesis Report. Accessed from: https://www.accesstoseeds.org/media/publications/.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, D. and Zilibotti, F. (2001). Productivity differences. *The Quarterly Journal of Economics*, 116(2):563–606.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–97.

Adamopoulos, T. and Restuccia, D. (2018). Geography and agricultural productivity: Cross-country evidence from micro plot-level data. Technical report, National Bureau of Economic Research.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Atkinson, A. B. and Stiglitz, J. E. (1969). A new view of technological change. *The Economic Journal*, 79(315):573–578.

Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in northern Mozambique. *The economic journal*, 116(514):869–902.

Barro, R. and Sala-i Martin, X. (1997). Technological diffusion, convergence, and growth. *Journal of Economic Growth*, 2(1):1–26.

Basu, S. and Weil, D. N. (1998). Appropriate technology and growth. *The Quarterly Journal of Economics*, 113(4):1025–1054.

Bazzi, S., Gaduh, A., Rothenberg, A. D., and Wong, M. (2016). Skill transferability, migration, and development: Evidence from population resettlement in Indonesia. *American Economic Review*, 106(9):2658–98.

Bebber, D. P., Holmes, T., and Gurr, S. J. (2014a). The global spread of crop pests and pathogens. *Global Ecology and Biogeography*, 23(12):1398–1407.

Bebber, D. P., Holmes, T., Smith, D., and Gurr, S. J. (2014b). Economic and physical determinants of the global distributions of crop pests and pathogens. *New Phytologist*, 202(3):901–910.

Bebber, D. P., Ramotowski, M. A., and Gurr, S. J. (2013). Crop pests and pathogens move polewards in a warming world. *Nature climate change*, 3(11):985–988.

Beintema, N., Stads, G.-J., Fuglie, K., and Heisey, P. (2012). ASTI global assessment of agricultural R&D spending: developing countries accelerate investment. Technical report, International Food Policy Research Institute. doi: 10.2499/9780896298026.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50.

Bento, J. (1999). Perdas por insetos na agricultura. *Ação ambiental*, 4(2):19–21.

Bessin, R. (2019). Bt-corn: What it is and how it works. Entfact 130, University of Kentucky College of Agriculture, Food and Environment. https://entomology.ca.uky.edu/ef130.

Binswanger, H. and Pingali, P. (1988). Technological priorities for farming in sub-Saharan Africa. *The World Bank Research Observer*, 3(1):81–98.

Boroush, M. (2020). Research and development: U.S. trends and international comparisons. Science and Engineering Indicators Report NSB-2020-3, National Science Foundation (NSF).

Campagne, P., Capdevielle-Dulac, C., Pasquet, R., Cornell, S., Kruger, M., Silvain, J.-F., LeRü, B., and Van den Berg, J. (2017). Genetic hitchhiking and resistance evolution to transgenic Bt toxins: insights from the African stalk borer Busseola fusca (Noctuidae). *Heredity*, 118(4):330–339.

Caselli, F. (2005). Accounting for cross-country income differences. *Handbook of Economic Growth*, 1:679–741.

Caselli, F. and Coleman II, W. J. (2006). The world technology frontier. *American Economic Review*, 96(3):499–522.

Caselli, F. and Wilson, D. J. (2004). Importing technology. *Journal of Monetary Economics*, 51(1):1–32.

Collinge, D. B., editor (2016). *Plant Pathogen Resistance Biotechnology*. John Wiley & Sons, Ltd, Hoboken, NJ, USA.

Comin, D. and Hobijn, B. (2004). Cross-country technology adoption: making the theories face the facts. *Journal of Monetary Economics*, 51(1):39–83.

Comin, D. and Hobijn, B. (2010). An exploration of technology diffusion. *American Economic Review*, 100(5):2031–59.

Comin, D. and Mestieri, M. (2014). Technology diffusion: Measurement, causes, and consequences. In *Handbook of Economic Growth*, volume 2, pages 565–622. Elsevier.

Comin, D. and Mestieri, M. (2018). If technology has arrived everywhere, why has income diverged? *American Economic Journal: Macroeconomics*, 10(3):137–78.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*, 100(1):35–69.

Costinot, A. and Donaldson, D. (2012). Ricardo's theory of comparative advantage: old idea, new evidence. *American Economic Review: Papers & Proceedings*, 102(3):453–58.

Costinot, A., Donaldson, D., and Smith, C. (2016). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. *Journal of Political Economy*, 124(1):205–248.

David, P. A. (1966). The mechanization of reaping in the ante-bellum midwest. In Rosovsky, H., editor, *Industrialization in Two Systems*, pages 3–39. New York: Wiley.

Diamond, J. (1997). *Guns, Germs, and Steel: the Fates of Human Societies*. WW Norton & Company, New York.

Dong, O. X. and Ronald, P. C. (2019). Genetic engineering for disease resistance in plants: recent progress and future perspectives. *Plant Physiology*, pages 26–38.

Duflo, E., Kremer, M., and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *American Economic Review*, 101(6):2350–90.

Eaton, J. and Kortum, S. (1996). Trade in ideas: Patenting and productivity in the OECD. *Journal of International Economics*, 40(3-4):251–278.

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

Engerman, S. L. and Sokoloff, K. L. (2002). Factor endowments, inequality, and paths of development among new world economies. *Economía*, 3(1):41–109.

Evenson, R. E. (2005). Besting Malthus: the Green Revolution. *Proceedings of the American Philosophical Society*, 149(4):469–486.

Evenson, R. E. and Gollin, D. (2003a). Assessing the impact of the Green Revolution, 1960 to 2000. *Science*, 300(5620):758–762.

Evenson, R. E. and Gollin, D., editors (2003b). *Crop variety improvement and its effect on productivity*. CABI Publishing, Cambridge, MA, USA.

Fand, B. B., Nagrare, V., Gawande, S., Nagrale, D., Naikwadi, B., Deshmukh, V., Gokte-Narkhedkar, N., and Waghmare, V. (2019). Widespread infestation of pink bollworm, Pectinophora gossypiella on Bt cotton in Central India: a new threat and concerns for cotton production. *Phytoparasitica*, 47(3):313–325.

Foster, A. D. and Rosenzweig, M. R. (1996). Technical change and human-capital returns and investments: evidence from the Green Revolution. *American Economic Review*, pages 931–953.

Foster, A. D. and Rosenzweig, M. R. (2004). Agricultural productivity growth, rural economic diversity, and economic reforms: India, 1970–2000. *Economic Development and Cultural Change*, 52(3):509–542.

Fuglie, K. (2015). Accounting for growth in global agriculture. *Bio-based and Applied Economics*, 4(3):201–234.

Fuglie, K. (2016). The growing role of the private sector in agricultural research and development world-wide. *Global Food Security*, 10:29–38.

Fuglie, K. O. (2012). Productivity growth and technology capital in the global agricultural economy. In Fuglie, K. O., Wang, S. L., and Ball, E., editors, *Productivity Growth in Agriculture: An International Perspective*, pages 335–368. CABI, Wallingford, UK.

Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J. R., Pangborn, W., and English, L. (2001). Structure of the insecticidal bacterial $\delta$-endotoxin Cry3Bb1 of Bacillus thuringiensis. *Acta Crystallographica Section D: Biological Crystallography*, 57(8):1101–1109.

Gallo, D., Nakano, O., Silveira Neto, S., Carvalho, R. L., Batista, G. d., Berti Filho, E., Parra, J. P., Zucchi, R., Alves, S., and Vendramim, J. (1988). *Manual de entomologia agrícola*. Agronômica Ceres São Paulo.

Gallup, J. L., Sachs, J. D., and Mellinger, A. D. (1999). Geography and economic development. *International Regional Science Review*, 22(2):179–232.

Giorcelli, M. (2019). The long-term effects of management and technology transfers. *American Economic Review*, 109(1):121–52.

Giorcelli, M. and Li, B. (2021). Technology transfer and early industrial development: Evidence from the sino-soviet alliance. *Available at SSRN*.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.

Gray, M. E., Sappington, T. W., Miller, N. J., Moeser, J., and Bohn, M. O. (2009). Adaptation and invasiveness of western corn rootworm: intensifying research on a worsening pest. *Annual Review of Entomology*, 54:303–321.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25(4):501–522.

Hardy, M. and McCasland, J. (2021). It takes two: Experimental evidence on the determinants of technology diffusion. *Journal of Development Economics*, 149:102600.

Herrera-Estrella, L. and Alvarez-Morales, A. (2001). Genetically modified crops: hope for developing countries? *EMBO Reports*, 2(4):256–258.

Hotez, P. J., Fenwick, A., Savioli, L., and Molyneux, D. H. (2009). Rescuing the bottom billion through control of neglected tropical diseases. *The Lancet*, 373(9674):1570–1575.

Hotez, P. J., Molyneux, D. H., Fenwick, A., Kumaresan, J., Sachs, S. E., Sachs, J. D., and Savioli, L. (2007). Control of neglected tropical diseases. *New England Journal of Medicine*, 357(10):1018–1027.

Jaccard, P. (1900). Contribution an probléme de l'immingration post-glaciaire de la flore alpine. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 36:87–130.

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:547–579.

Jerzmanowski, M. (2007). Total factor productivity differences: Appropriate technology vs. efficiency. *European Economic Review*, 51(8):2080–2110.

Jost, L., Chao, A., and Chazdon, R. L. (2011). Compositional similarity and $\beta$ (beta) diversity. In Magurran, A. E. and McGill, B. J., editors, *Biological Diversity: Frontiers in Measurement and Assessment*, pages 66–84. Oxford University Press, New York.

Kamarck, A. M. (1976). *The tropics and economic development*. The John Hopkins University Press, Baltimore, MD, USA.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

Kerr, W. R. (2008). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics*, 90(3):518–537.

Kloppenburg, J. R. (2005). *First the seed: The political economy of plant biotechnology*. Univ of Wisconsin Press.

Kremer, M. (2002). Pharmaceuticals and the developing world. *Journal of Economic Perspectives*, 16(4):67–90.

Kremer, M. and Glennerster, R. (2004). *Strong medicine: creating incentives for pharmaceutical research on neglected diseases*. Princeton University Press.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959.

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–80.

Lansing, J. S. (2009). *Priests and programmers: technologies of power in the engineered landscape of Bali*. Princeton University Press.

McMullen, N. (1987). Seeds and world agricultural progress. Technical report, National Planning Association.

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

Moseman, A. H. (1970). *Building agricultural research systems in the developing countries.* Agricultural Development Council, New York.

Muhammad, A., Seale, J. L., Meade, B., and Regmi, A. (2011). International evidence on food consumption patterns: an update using 2005 international comparison program data. Technical Bulletin 1929, USDA-ERS.

Munkvold, G. P. and Hellmich, R. L. (1999). Genetically modified insect resistant corn: Implications for disease management. *APSnet Plant Pathology On-line Feature*, 15.

Nordhaus, H. (2017). Cornboy vs. the billion-dollar bug (cover story). *Scientific American*, 316(3):64–71.

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

Oerke, E.-C. and Dehne, H.-W. (2004). Safeguarding production—losses in major crops and the role of crop protection. *Crop Protection*, 23(4):275–285.

Olmstead, A. L. and Rhode, P. W. (2008). *Creating Abundance: Biological Innovation and American Agricultural Development*. Cambridge University Press, New York.

Paini, D. R., Sheppard, A. W., Cook, D. C., De Barro, P. J., Worner, S. P., and Thomas, M. B. (2016). Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences*, 113(27):7575–7579.

Parente, S. L. and Prescott, E. C. (1994). Barriers to technology adoption and development. *Journal of Political Economy*, 102(2):298–321.

Parente, S. L. and Prescott, E. C. (2002). *Barriers to Riches*. MIT Press.

Parmesan, C. and Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918):37–42.

Pasiecznik, N., Smith, I., Watson, G., Brunt, A., Ritchie, B., and Charles, L. (2005). CABI/EPPO distribution maps of plant pests and plant diseases and their important role in plant quarantine. *Eppo Bulletin*, 35(1):1–7.

Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences*, 109(31):12302–12308.

Reynolds, M. P. and Borlaug, N. (2006). Impacts of breeding on international collaborative wheat improvement. *The Journal of Agricultural Science*, 144(1):3–17.

Ruttan, V. W. and Hayami, Y. (1973). Technology transfer and agricultural development. *Technology and Culture*, 14(2):119–151.

Sachs, J. D. and Warner, A. M. (1997). Sources of Slow Growth in African Economies. *Journal of African Economies*, 6(3):335–376.

Salter, W. E. G. (1969). *Productivity and Technical Change*, volume 960. Cambridge University Press.

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology & evolution*, 3(3):430–439.

Sayre, K., Singh, R., Huerta-Espino, J., and Rajaram, S. (1998). Genetic progress in reducing losses to leaf rust in CIMMYT-derived Mexican spring wheat cultivars. *Crop Science*, 38(3):654–659.

Shaw, M. W. and Osborne, T. M. (2011). Geographic distribution of plant pathogens in response to climate change. *Plant Pathology*, 60(1):31–43.

Sokoloff, K. L. and Engerman, S. L. (2000). Institutions, factor endowments, and paths of development in the new world. *Journal of Economic perspectives*, 14(3):217–232.

Sotelo, S. (2020). Domestic trade frictions and agriculture. *Journal of Political Economy*, 128(7):2690–2738.

Stewart, F. (1978). *Technology and Underdevelopment*. MacMillan, London, UK.

Stone, M. (2020). A plague of locusts has descended on East Africa. Climate change may be to blame. *National Geographic: Science*. Retrieved from: https://www.nationalgeographic.com/science/article/locust-plague-climate-science-east-africa.

Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica*, 79(1):159–209.

Tabashnik, B. E. and Carrière, Y. (2019). Global patterns of resistance to Bt crops highlighting pink bollworm in the United States, China, and India. *Journal of Economic Entomology*, 112(6):2513–2523.

Union for the Protection of New Varieties of Plants (1991). International convention for the protection of new varieties of plants. First adopted on december 2, 1961. UPOV Publication no. 221(e). https://upovlex.upov.int/en/convention.

Union for the Protection of New Varieties of Plants (2015). Explanatory notes on variety denominations under the upov convention. Adopted by the upov council on october 29, 2015. UPOV Document UPOV/INF/12/5. https://www.upov.int/edocs/infdocs/en/upov_inf_12.pdf.

Van Esse, H. P., Reuber, T. L., and van der Does, D. (2020). Genetic modification to improve disease resistance in crops. *New Phytologist*, 225(1):70–86.

Vanderplank, J. E. (2012). *Disease resistance in plants*. Academic Press, Orlando, FL, USA.

Verhoogen, E. (2021). Firm-level upgrading in developing countries.

Vidal, J. (2014). Gates foundation spends bulk of agriculture grants in rich countries. *The Guardian*. November 3. Retrieved from: https://www.theguardian.com/global-development/2014/nov/04/bill-melinda-gates-foundation-grants-usa-uk-africa.

Waage, J. and Mumford, J. D. (2008). Agricultural biosecurity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1492):863–876.

**Online Appendix** for

for "Inappropriate Technology: Evidence from Global Agriculture"

by Moscona and Sastry

## Contents



# A. Omitted Proofs and Derivations

## A.1 Statement and Proof of Lemma 1

We first state and prove a result deriving the optimal planting patterns in each country.

**Lemma 1.** *The measure of farmers planting crop $k$ with technology $\ell'$ in country $\ell$ is given by*

$$\pi(k,\ell',\ell) = \frac{p(k)^\eta \theta(k,\ell' \to \ell)^\eta \omega(k,\ell)^\eta}{\sum_{k',\ell''} p(k')^\eta \theta(k',\ell'' \to \ell)^\eta \omega(k',\ell)^\eta} \tag{A.1}$$

*Proof.* Let $u_i^* \in \{1,\ldots,K\} \times \{1,\ldots,L\}$ denote the crop-technology choice of farmer $i$, let $\nu(k,\ell',\ell) = p(k)\omega(k,\ell)\theta(k,\ell' \to \ell)$ determine the shifters of revenue productivity for each $(k,\ell')$ pair in $\ell$, and let $\pi(k,\ell',\ell) = \mathbb{P}[u_i^* = (\ell',k)]$ if $i \in [\ell-1,\ell)$, which does not depend on the index $i$ within a given farming country $\ell$.[62] Let $F(z)$ denote the cumulative distribution function of a Fréchet random variable with

[62]By a law of large numbers across i.i.d. realizations of the shocks, this corresponds with the measure of farmers making the specified choice.

scale one and shape parameter $\eta > 1$, or

$$F(z) = \exp\left(-x^{\eta}\right) \tag{A.2}$$

The random shock $\varepsilon_i(i,\ell')$ is Fréchet random variable with *mean* one and shape parameter $\eta > 1$, so its scale parameter is $s = (\Gamma(1-1/\eta))^{-1}$; thus the normalized shock $\hat{\varepsilon}_i(i,\ell') = \frac{1}{s}\varepsilon_i(i,\ell')$ is distributed by $F(z)$. If a farmer draws $\hat{\varepsilon}_i(k,\ell') = z$ for their random productivity, then that farmer chooses pair $(k,\ell')$ only if this results in the maximum productivity among all options, or

$$\nu(k,\ell',\ell)z > \nu(k',\ell'',\ell)\hat{\varepsilon}_i(k',\ell'') \tag{A.3}$$

for all other pairs $(k',\ell'')$. These events are independent across all $(k',\ell'')$. Thus the probability of choosing $(k,\ell')$ is given by the probability of the event described above, conditional on each realization $z$, integrated over the probability distribution of $z$:

$$\pi(k,\ell' \to \ell) = \int_0^{\infty} \prod_{k',\ell''\neq k,\ell'} F\left(\frac{\nu(k,\ell',\ell)}{\nu(k',\ell'',\ell)}z\right) \mathrm{d}F(z) \tag{A.4}$$

Substituting Equation A.2 into Equation A.4 and simplifying yields the expression

$$\pi(k,\ell',\ell) = \int_0^{\infty} \exp\left(-z^{-\eta}\frac{\Xi(\ell)^{\eta}}{\nu(k,\ell',\ell)^{\eta}}\right) z^{-1-\eta}\,\mathrm{d}z \tag{A.5}$$

where we define the productivity index

$$\Xi(\ell) = \left(\sum_{k=1}^{K}\sum_{\ell'=1}^{L}\nu(k,\ell',\ell)^{\eta}\right)^{\frac{1}{\eta}} \tag{A.6}$$

which corresponds with the index defined in Equation 2.6 in the main text. See that, after a change in variables in the integrand to

$$\tilde{z} = z\frac{\nu(k,\ell',\ell)}{\Phi(\ell)} \tag{A.7}$$

that the original integral can be re-written and simplified as

$$\begin{aligned}
\pi(k,\ell',\ell) &= \frac{\nu(k,\ell',\ell)^{\eta}}{\sum_{k',\ell''}\nu(k',\ell'',\ell)^{\eta}}\int_0^{\infty}\exp\left(-\tilde{z}^{-\eta}\right)\tilde{z}^{-1-\eta}\,\mathrm{d}z \\
&= \frac{\nu(k,\ell',\ell)^{\eta}}{\sum_{k',\ell''}\nu(k',\ell'',\ell)^{\eta}}\int_0^{\infty}\mathrm{d}F(\tilde{z}) \\
&= \frac{\nu(k,\ell',\ell)^{\eta}}{\sum_{k',\ell''}\nu(k',\ell'',\ell)^{\eta}}
\end{aligned} \tag{A.8}$$

Re-writing the last line with the definition of $\nu(k,\ell',\ell)$ completes the proof □

## A.2 Proof of Proposition 1

We first derive the program of an innovator for profit-maximizing research. Fix the innovator's country $\ell'$ and crop $k$. Since individual innovators are small, the measure of farmers using any technology $j \in [\ell' - 1, \ell')$ in country $\ell$ for crop $k$ is given by $\pi(k, \ell', \ell)$ as derived in Lemma 1 and Equation A.1. If all innovators have the same technology quality, or $\theta_j(k, \ell') \equiv \theta(k, \ell' \to \ell)$, then the productivity of each farmer planting $(k, \ell')$ is given by $\Xi(\ell)$, defined in Equation in 2.6, due to the arguments in the proof of Proposition 2. Finally, the productivity of a farmer planting technology $j$ is given by $\Xi(\ell)\frac{\theta_j(k,\ell)}{\theta(k,\ell'\to\ell)}$, where the second term measures any differential productivity relative to the average quality of technology. The innovator chooses general research, or $A_j(k)$, and pest-specific research, or the mapping $t, \ell \mapsto B_j(t, k, \ell)$, to solve the program

$$\max_{A_j(k), B_j(t,k,\ell)>0} \sum_{\ell=1}^{L} \rho(\ell, \ell')\pi(k, \ell', \ell)\Xi(\ell)\frac{\theta_j(t, k, \ell')}{\theta(k, \ell' \to \ell)} - \sum_{\ell=1}^{L} \int_{\mathcal{T}} C(B_j(t, k, \ell); t, k, \ell')\,\mathrm{d}t - Q(A_j(k)) \quad \text{(A.9)}$$

where $\mathcal{T}$ denotes the set of all pests and $Q(\cdot)$ denotes the cost of researching general technology, which we assume to be convex. The program is concave, owing to the concavity of the objective (which is "Cobb-Douglas," with constant returns to scale) and convexity of all costs, and its solution is characterized by necessary first-order conditions for each choice variable.

We first establish two basic observations about pest-specific research within the home country $\ell'$. See that, for any $k$ and $t \notin \mathcal{T}(k, \ell')$, there is zero marginal benefit to research. Therefore, it is necessary in any optimal allocation for $B_j(t, k, \ell') \equiv 0$ for all $k$ and $t \notin \mathcal{T}(k, \ell')$. Next, see that the first-order condition for any $k$ and $t \in \mathcal{T}(k, \ell)$ is

$$(1 - \alpha)\rho(\ell', \ell')\pi(k, \ell', \ell')\Xi(\ell')\frac{\theta_j(t, k, \ell')}{\theta(k, \ell' \to \ell')} = B_0^{\phi+1} B_j(t, k, \ell')^{\phi+1} \exp(-\tau(B(t, k, \ell'))) \quad \text{(A.10)}$$

Under a symmetric equilibrium, this has a unique solution $B(k, \ell') > 0$ for any specific pest.

We next focus on the first-order conditions for each $B_j(t, k, \ell)$ for $\ell \neq \ell$. There are three cases. First, $t \notin \mathcal{T}(k, \ell)$ or the pest is not present, marginal benefits are zero and optimal investment is zero. Second, if $t \in \mathcal{T}(k, \ell)$ and $t \notin \mathcal{T}(k, \ell')$, then the first-order condition is given by the following

$$(1 - \alpha)\rho(\ell, \ell')\pi(k, \ell', \ell)\Xi(\ell)\frac{\theta_j(t, k, \ell')}{\theta(k, \ell' \to \ell)} = B_0^{\phi+1} B_j(t, k, \ell)^{\phi+1} \quad \text{(A.11)}$$

incorporating the zero knowledge spillover, from zero $\ell'$ research. Finally, if $t \in \mathcal{T}(k, \ell)$ and $t \in \mathcal{T}(k, \ell')$, then the first-order condition is given by the following that incorporates the knowledge spillover:

$$(1 - \alpha)\rho(\ell, \ell')\pi(k, \ell', \ell)\Xi(\ell)\frac{\theta_j(t, k, \ell')}{\theta(k, \ell' \to \ell)} = B_0^{\phi+1} B_j(t, k, \ell)^{\phi+1} \exp(-\tau(B(k, \ell'))) \quad \text{(A.12)}$$

We focus on symmetric equilibria in which $\theta_j(t,k,\ell') \equiv \theta(k,\ell' \to \ell)$ for all $j \in [\ell'-1,\ell)$ and $B(t,k,\ell')^\tau \equiv B(k,\ell')^\tau$ for all $t \in \mathcal{T}(k,\ell')$. In this case, $\frac{\theta_j(t,k,\ell')}{\theta(k,\ell' \to \ell)} = 1$ in each equation.

We now derive the the expression for technology transfer, Equation 2.4. Taking logs, integrating Equations A.11 and A.12 over all pests $t \in \mathcal{T}(k,\ell)$, and adding $\log A(k,\ell')$, we derive the following condition for $\log \theta(k,\ell' \to \ell)$:

$$\frac{1+\phi}{1-\alpha}\log\theta(k,\ell'\to\ell) = \log(1-\alpha) - (1+\phi)\log B_0 + \log\rho(\ell,\ell') + \tau B(k,\ell') - \delta(k,\ell',\ell)\tau(B(k,\ell')) + \log\Xi(\ell) + \log\pi(k,\ell',\ell) + \frac{\alpha}{1-\alpha}(1+\phi)\log A(k,\ell') \tag{A.13}$$

Substituting in the expression for $\pi(k,\ell',\ell)$ from Lemma 1, this re-arranges as desired to

$$\log\theta(k,\ell'\to\ell) = \beta(k,\ell')\cdot\delta(k,\ell',\ell) + \chi(k,\ell) + \chi(k,\ell') + \chi(\ell,\ell') \tag{A.14}$$

with the fixed effects defined by

$$\begin{aligned}
\chi(k,\ell) &= \frac{1-\alpha}{1+\phi-(1-\alpha)\eta}\left(\eta\log p(k) + \eta\log\omega(k,\ell) - (\eta-1)\Xi(\ell)\right)\\
\chi(k,\ell') &= \frac{1-\alpha}{1+\phi-(1-\alpha)\eta}\left(\tau(B(k,\ell')) + (1+\phi)\log A(k,\ell')\right)\\
\chi(\ell',\ell) &= \frac{(1-\alpha)(\log\rho(\ell',\ell) + \log(1-\alpha) - (1+\phi)\log B_0)}{1+\phi-(1-\alpha)\eta}
\end{aligned} \tag{A.15}$$

and coefficient

$$\beta(k,\ell') = -\frac{(1-\alpha)\tau(B(k,\ell'))}{1+\phi-(1-\alpha)\eta} \tag{A.16}$$

As $1+\phi-(1-\alpha)\eta > 0$, by the assumption stated in Footnote 16, we furthermore have that $\beta(k,\ell') \le 0$.

### A.3 Proof of Proposition 2

We first derive productivity of $k,\ell'$ production conditional on choice. Let

$$V_i^* = \max_{k',\ell''}\{\psi_i(k',\ell'')\} \tag{A.17}$$

denote the productivity of farmer $i$ evaluated at the optimal choice. The probability that $V_i^*$ is less than some value $v$, conditional on the optimal choice being $(k',\ell'')$, can be obtained by integrating the right-hand-side of Equation A.4 up to the realization $\frac{v}{s\nu(k',\ell'',\ell)^\eta}$, and normalizing by the probability of choosing $(k',\ell'')$:

$$\mathbb{P}[V_i^* \le v \mid u_i^* = (k',\ell'')] = \frac{1}{\pi(k,\ell'\to\ell)}\int_0^{\frac{v}{s\nu(k',\ell'',\ell)^\eta}} F\left(\frac{\nu(k,\ell',\ell)^\eta}{\nu(k',\ell'',\ell)^\eta}z\right)\mathrm{d}F(z) \tag{A.18}$$

where $\nu(k, \ell', \ell) = p(k)\omega(k, \ell)\theta(k, \ell' \to \ell)$ as defined previously. Doing the same manipulation of the integrand and change-of-variables as in the proof of Lemma 1, we can re-express this probability as

$$\mathbb{P}[V_i^* \leq v \mid u_i^* = (k', \ell'')] = \int_0^{\frac{v}{s\Xi(\ell)}} \mathrm{d}F(\tilde{z}) \tag{A.19}$$

which implies that $V_i^*$, conditional on $u_i^* = (k', \ell'')$, can be written as the product of $\Xi(\ell)$ and a unit-mean, $\eta$-shape Fréchet random variable. Since this is invariant to $k', \ell''$, this is also the unconditional distribution of $V_i^*$. Moreover, it implies that $\mathbb{E}[V_i^* | u_i^* = (k', \ell'')] = \Xi(\ell)$ for any $(k', \ell'')$, as well as unconditionally.

We first derive an expression for the physical yield of crop $k$ in country $\ell$. Because of the law of large numbers, this is equal to the expected physical production per unit area:

$$z(k, \ell) = \frac{1}{p(k)} \mathbb{E}\left[V_i^* \mid u_i^* = (k, \ell'), \ell' \in \{1, \ldots, L\}\right] \tag{A.20}$$

As established above, the conditional expectation is $\Xi(\ell)$. Thus, $z(k, \ell) = \frac{\Xi(\ell)}{p(k)}$.

Next, see that the planted area equals the probability of selecting crop $k$ within any location $\ell'$, again due to the law of large numbers over individual farms. This probability is

$$\pi(k, \ell) = \frac{\sum_{\ell'}^{L} \nu(k, \ell', \ell)^{\eta}}{\sum_{k', \ell''} \nu(k', \ell'', \ell)^{\eta}} \tag{A.21}$$

Combining the previous with the definitions of $\Xi(\ell)$ and $\nu(k, \ell', \ell)$, and taking a log, we derive the following expression in terms of primitives:

$$\log \pi(k, \ell) = \eta \log \Theta(k, \ell) + \eta \log \omega(k, \ell) + \eta \log p(k) - \eta \log \Xi(\ell) \tag{A.22}$$

Finally, see that physical production can be written as

$$Y(k, \ell) = \sum_{\ell'}^{L} \mathbb{E}\left[\frac{V_i^*}{p(k)} \mid u_i^* = (k, \ell')\right] \cdot \pi(k, \ell', \ell) \tag{A.23}$$

By the arguments above, $\mathbb{E}\left[\frac{V_i^*}{p(k)} \mid u_i^* = (k, \ell')\right] \equiv \frac{\Xi(\ell)}{p(k)}$, or the uniform physical yield, and hence

$$Y(k, \ell) = \frac{\Xi(\ell)\pi(k, \ell)}{p(k)} \tag{A.24}$$

Combining this with Equation A.22, and taking a log, yields

$$\log Y(k, \ell) = \eta \log \Theta(k, \ell) + \eta \log \omega(k, \ell) + (\eta - 1) \log p(k) + (1 - \eta) \log \Xi(\ell) \tag{A.25}$$

This proves the claim of Proposition 2.

### A.4 Statement and Proof of Corollary 1

**Corollary 1.** *The fraction of crop k farmers using technology from country $\ell'$ in location $\ell$ is given by*

$$\log \pi(k, \ell' \to \ell) = \eta \cdot \beta(k, \ell') \cdot \delta(k, \ell', \ell) + \hat{\chi}(k, \ell) + \hat{\chi}(k, \ell') + \hat{\chi}(\ell, \ell') \tag{A.26}$$

*where $\beta(k, \ell') \leq 0$ is given in Equation 2.5, and the $\hat{\chi}(\cdot)$ are additive effects varying at the indicated level.*

First, see that we can write the conditional probability of using technology from $\ell'$ in terms of the probabilities of choosing each $(k, \ell')$ pair:

$$\pi(\ell' \mid k, \ell) = \frac{\pi(k, \ell', \ell)}{\sum_{\ell''=1}^{L} \pi(k, \ell'', \ell)} \tag{A.27}$$

Applying Lemma 1, and simplifying, we derive

$$\begin{aligned} \pi(\ell' \mid k, \ell) &= \frac{p(k)^{\eta} \theta(k, \ell' \to \ell)^{\eta} \omega(k, \ell')^{\eta}}{\sum_{\ell''=1}^{L} p(k)^{\eta} \theta(k, \ell'' \to \ell)^{\eta} \omega(k, \ell)^{\eta}} \\ &= \frac{\theta(k, \ell' \to \ell)^{\eta}}{\sum_{\ell''=1}^{L} \theta(k, \ell'' \to \ell)^{\eta}} \end{aligned} \tag{A.28}$$

We finally take logs to derive Equation A.26, defining the fixed effects as

$$\begin{aligned} \hat{\chi}(k, \ell) &= \eta \chi(k, \ell) - \log \left( \sum_{\ell''=1}^{L} \theta(k, \ell'' \to \ell)^{\eta} \right) \\ \hat{\chi}(k, \ell') &= \eta \chi(k, \ell') \\ \hat{\chi}(\ell', \ell) &= \eta \chi(\ell', \ell) \end{aligned} \tag{A.29}$$

where $(\chi(k, \ell), \chi(k, \ell'), \chi(\ell', \ell))$ are as in Equation 2.4, and as defined in the proof of Proposition 1.

# B. Additional Empirical Analysis

## B.1 Invasive Species

In our baseline estimates, we construct CPP mismatch using all known pests and pathogens present in each country that affect each crop. This measure captures the true extent of global differences in CPP ecology across crops and countries. An important conceptual question is whether the baseline findings are driven by invasive species, or persistent differences in ecology across crops and locations. Invasive species can cause disproportionate damage to plants and agricultural production since they often have fewer natural predators in the new environment, and other species have not evolved natural defense mechanisms. Moreover, if the results are strongly driven by invasive species, it would be important to explore further the causes of species movement and ensure that they are not correlated with omitted factors that could drive our results. However, as discussed in the main text, there are several examples of persistent differences in CPP environment across locations shaping the effectiveness of technology (see Section 3.1).

To investigate the role of invasive species, we use an additional data set produced by CABI: the Invasive Species Compendium (ISC).[63] The ISC is a list of global invasive species, as determined by extensive literature searches and trawls of existing invasive species lists. Since the ISC is also a CABI data set, we can use the unique species identifiers to link ISC species to CPC species in our main CPP data set. 748 CPPs from our main sample are listed as invasive species in the ISC, comprising about 15% of our main CPP sample. We then estimate all versions of CPP distance from the main text *after restricting the sample of CPPs to non-invasive species*, and re-produce all of our main estimates using the CPP distance measures purged of variation from invasive species.[64]

The estimates are presented in Table A12. Columns 1-3 report estimates corresponding to our analysis of international technology diffusion and columns 4 and 5 report estimates correspond to our analysis of biotechnology adoption and output respectively. Compared to our baseline estimates, the effects on technology diffusion are (if anything) slightly larger, and the effects on output are slightly smaller (although the standardized effect is similar, since the standard deviation of CPP distance without invasive species is somewhat smaller). These findings suggest that the baseline results are not driven by invasive species.

## B.2 Inappropriateness Driven By Non-CPP Agro-Climatic Conditions

This section investigates the possible importance of non-CPP agro-climatic conditions as shifters of ecological inappropriateness. We estimate ecological differences across crop-specific growing areas in different countries and study how these differences shape technology diffusion and crop-level output.

[63]The ISC homepage can be found here: https://www.cabi.org/isc

[64]It could be ideal to only exclude country-CPP pairs where the CPP is known to by invasive. However, we are unaware of systematic data on the locations of species invasion; CABI do not report this level of detail.

We also investigate the relationship between these measures of geographic mismatch and our baseline CPP-derived measure. Finally, we estimate our baseline counterfactual scenario incorporating these non-CPP differences in crop-specific growing conditions.

### B.2.1 Constructing Agro-climatic Mismatch

We include ten key agroclimatic characteristics that shape the usefulness of biotechnology for production in a region: temperature, precipitation, elevation, ruggedness, the growing season, and soil acidity, clay content, silt content, coarse fragment content, and water capacity.[65] We combine geographically coded raster files of each aforementioned characteristic with grid-cell level information from the EarthStat database, which reports the global planting pattern of 175 important crops in the year 2000 (Monfreda et al., 2008).[66] We then compute the value of each characteristic for each *crop-by-country pair* by estimating the average value of each characteristic in each country *on just the land that EarthStat identifies is devoted to the crop in question*; we denote these as $x_{k,\ell}$. We then simply normalize each characteristic so that all are in comparable units by re-centering by the global mean value of each attribute and normalizing by the global dispersion (standard deviation); we refer to these normalized values as $\hat{x}_{k,\ell}$. For each agro-climatic characteristic $x \in \mathcal{X}$ we define:

$$\Delta \hat{x}_{k,\ell,\ell'} = |\hat{x}_{k,\ell} - \hat{x}_{k,\ell'}| \tag{B.1}$$

where, in words, $\hat{x}_{k,\ell,\ell'}$ is the normalized distance ("inappropriateness") in agro-climatic feature $x$ for crop $k$ between countries $\ell$ and $\ell'$. For simplicity, we also aggregate the individual agroclimatic characteristics into a single index at the crop-by-country-pair level:

$$\text{AgroClimMismatch}_{k,\ell,\ell'} = \frac{1}{|\mathcal{X}|} \cdot \sum_{x \in |\mathcal{X}|} |\hat{x}_{k,\ell} - \hat{x}_{k,\ell'}| \tag{B.2}$$

where $\mathcal{X}$ is the set of agro-climatic characteristics $x$. The index is similar to the agro-climatic similarity index used by Bazzi et al. (2016) to study patterns of migration. This index has the attractive feature that it is additively separable the $x$'s and therefore simple to separate the contribution of each attribute.

### B.2.2 Empirical Estimates

We next investigate the role of differences across agro-climatic features in shaping the transfer of technology and productivity differences. Column 1 of Table A13 re-produces our baseline estimate of Equation 4.1 on the sample of country-pairs and crops for which all agro-climatic features could be

[65] This set of characteristics builds from earlier work on the transferability of agricultural knowledge across ecologically different regions (see, for example, Bazzi et al., 2016).

[66] The data set is described and can be accessed here: http://www.earthstat.org/harvested-area-yield-175-crops/. The data set was created by combining national, state, and county level census data with information on crop-specific maximum potential yield around the world, to construct a 5-minute by 5-minute grid of the area devoted to each of 175 important crops circa the year 2000.

measured. Our estimate is negative, significant, and slightly larger in magnitude than our estimate on the largest possible sample reported in the main text.

In column 2, we add $\Delta x_{k,\ell,\ell'}$ for all $x \in \mathcal{X}$. Consistent with agricultural biotechnology also being specific to particular non-CPP features of the environment (e.g. via repeated selection in a particular local environment), the coefficients on the $\Delta x_{k,\ell,\ell'}$ are almost all negative and some are statistically significant. Mismatch in temperature and precipitation are associated with the largest reductions in technology transfer, and there is also a significant effect of mismatch in elevation and soil pH. Despite the inclusion of all these additional distance metrics, however, the coefficient on CPP mismatch barely changes. In column 3, we include the one-dimensional $\text{AgroClimMismatch}_{k,\ell,\ell'}$ on the right hand side of the regression in place of the individual characteristics. The coefficient on agro-climatic mismatch is negative and significant; however, the coefficient on CPP distance again remains very similar, suggesting that non-CPP ecological differences do not bias our baseline estimates.

In Table A14, the dependent variable is log of agricultural output and the regression specification is Equation 5.2. Column 1 reproduces our main result, the relationship between CPP mismatch with the frontier and output, on the reduced sample on which we were able to estimates all agro-climatic characteristics. The specification in column 2 includes both CPP mismatch and agro-climatic mismatch on the right hand side. While mismatch with the frontier in non-CPP agro-climatic features significantly lower output, these effects again operate largely independently from CPP mismatch.

Taken together, these results show that our main findings are not specific to CPP differences across crops and places (or, more perniciously, not driven by some specific feature of our CPP data and measurement strategy); other agro-climatic shifters of inappropriateness also affect technology transfer and productivity gaps. At the same time, non-CPP agro-climatic differences as we measure them seem to operate independently from our baseline measure of CPP mismatch, suggesting that the baseline estimates are not simply picking up standard features of climate and geography.

These findings are all consistent with the fact that the pairwise correlations between CPP mismatch with the frontier, and mismatch with the frontier in each other ecological characteristic, is relatively low. Table A15 reports a correlation matrix, including CPP distance to the frontier along with all agro-climatic characteristics discussed above. The first column shows the correlation between CPP distance and all other distance measures; the correlation coefficients tend to be small, and only one is above 0.2. Several are 0.1 or below. The remainder of the table includes correlation coefficients among all other pairs of ecological characteristics.

Finally, we estimate our baseline counterfactuals scenario incorporating both CPP mismatch and acro-climatic mismatch, using the estimates from column 3 of Table A14. Our empirical strategy is identical to the one outlined in Section 7.1 of the main text. We find that inappropriateness, as captured by both CPP mismatch and agro-climatic mismatch, reduces global productivity by 52% and increases disparities in global productivity across countries by 16%. These results are summarized graphically in Figure A10, which is structured in the same way as Figure 5 in the main text. Thus, incorporating agro-climatic mismatch as an additional shifter of inappropriateness increases our estimate of the overall effect of inappropriateness on productivity. However, as foreshadowed by the reduced form estimates in Table A14, the effect of CPP mismatch on global output is about four times as large as the effect of agro-climatic mismatch, suggesting that inappropriateness in the form of agricultural pests and pathogens is a particularly important determinant of global agricultural productivity.

### B.3 The Global Direction of Agricultural Innovation

The inappropriate technology hypothesis is based on the premise that global innovation is biased toward the needs and demands of wealthy frontier countries. There are two reasons we expect this bias to exist, which were both implicit throughout the examples given so far. First, if innovation is more likely to occur in rich countries with more biotechnological infrastructure, it may take advantage of local "technology production opportunities." This mechanism is embodied in the local knowledge spillovers in the model, and may take the general form of accumulated expertise, available test fields for breeding or trials, and readily available germplasm for genetic analysis. Second, since wealthy countries tend to be large markets, global innovation which occurs anywhere in the world may still be directed toward their needs as part of profit-maximizing behavior.

We explore both of these hypotheses in reduced form within our global varieties data (UPOV PLUTO), focusing on novel plant varieties released anywhere in the world in the 21st century. Let $\text{BioTech}_k$ be the count of all unique denominations produced in the world for crop $k$ over this period; this will be our simple measure of global technological progress for a given crop. To quantify the targeting of this technology, measured in this simple way, we estimate the following regression model:

$$\log(\text{BioTech}_k) = \alpha + \delta_1 \cdot \log \text{WorldArea}_k + \delta_2 \cdot \log \text{GDPArea}_k + \delta_3 \cdot \log \text{IPArea}_k + \varepsilon_k \tag{B.3}$$

in which $\log \text{WorldArea}_k$ is the (log of) global area devoted to crop $k$, and the other two regressors are respectively this area weighted by per-capita GDP (averaged over 1990-1999) and the presence of intellectual property for plant varieties as of 2000:[67]

$$\log \text{GDPArea}_k = \log\Big(\sum_{\ell} \text{Area}_{k,\ell} \cdot \text{GDP}_{\ell}\Big) \qquad \log \text{IPArea}_k = \log\Big(\sum_{\ell} \text{Area}_{k,\ell} \cdot \mathbb{I}_{\ell}^{IP}\Big) \tag{B.4}$$

We think of the first regressor, and its coefficient $\delta_1$, as (to first approximation) a proxy for each crops importance to global livelihoods when *not* adjusted by production and/or willingness to pay for technology; while the latter two regressors, and their coefficients $(\delta_2, \delta_3)$, could each capture bias via either channel described above.

Figure A6 reports our estimates of $\delta_2$ and $\delta_3$, in the form of partial correlation plots in which each dot is a crop. Consistent with the hypothesis, both are positive and significant, and together have

[67] We compile the latter data using UPOV's collation of relevant intellectual property law across its member states, under the premise that participation in UPOV is essentially universal conditional on having meaningful IP protection.

an incremental $R^2$ of 29%. To give a sense of the estimated magnitudes, suppose the global market size of cotton increased by 1%; the regression estimates imply that, if this expansion occurred in the United States, the number of cotton varieties developed would increase by 4.41%; if it occurred in Brazil, a less wealthy country but one that *does* protect intellectual property, the number of cotton varieties developed would increase by 1.31%; and if it occurred in India, a low-income country that does *not* protect intellectual property, there would be essentially no effect.

To offer reduced-form clues that can distinguish between the two possible causes of this bias described above, we also estimate the following model at the level of crop-$k$ and country-$\ell$ pairs:

$$\log(\text{BioTech}_{k,\ell}) = \delta_0 \cdot \log \text{Area}_{k,\ell} + \delta_1 \cdot \log \text{WorldArea}_k + \delta_2 \cdot \log \text{GDPArea}_k + \delta_3 \cdot \log \text{IPArea}_k + \chi_\ell + \varepsilon_{k,\ell} \tag{B.5}$$

in which $\text{BioTech}_{k,\ell}$ is the number of varieties of crop $k$ developed in country $\ell$ since 2000; and $\chi_\ell$ are country fixed effects. The term $\delta_0 \cdot \log \text{Area}_{k,\ell}$ isolates "local focus," potentially due to local specificity of technology production, relative to all innovators' uniform desire to cater to large markets, as captured by the next three terms. Estimates of Equation (B.5) are reported in Table A16. We find that $\delta_0 \gg 0$, suggesting that the local focus of innovators an important mechanism; $\delta_2$ and $\delta_3$ are also positive, although only marginally significant. Finally, in this framework, $\delta_1 = 0$; un-weighted global market size is uncorrelated with technology development.

Together, this evidence suggests that in our data, technology development is biased toward the demands of wealthy, IP-protecting countries; this effect appears driven by the fact that innovation takes place *in* these countries and innovators develop technology for their home markets. These estimates mirror our findings using the CPP-specific patent data in Section 3.3 and further motivate the local R&D spillovers in the model.

### B.4 Technology Transfer to Africa

The UPOV data set tracks all plant variety certificates and as a result only covers countries for which intellectual property protection is in place; this results, as can be seen in Figure A2, several omissions, most notably much of Africa. Moreover, despite the fact that barriers to obtaining protection under UPOV are low and the fact that public entities often obtain protection (e.g. the Sicot Cotton example from the main text), there is a potential concern that our main technology diffusion data under-samples public sector innovation. To partially fill this gap, especially in light of our subsequent analysis on technology adoption focuses on sub-Saharan Africa, we compile data from the Consultative Group on International Agricultural Research (CGIAR) Diffusion and Impact of Improved Varieties in Africa (DIIVA) project. DIIVA has collected data on improved crop varieties for 28 countries in sub-Saharan Africa and across 19 crops since 1960, and incorporates an extensive search of public-sector research and variety release.

Using the DIIVA Project data, we compute the number of varieties for each plant species introduced in 28 African countries; since we do not know the country of origin of each variety, in order to investigate whether inappropriateness is a barrier to technology using these data, we estimate a simplified version of (4.1):

$$y_{k,\ell} = \beta \cdot \text{CPPMismatch}_{k,\ell} + \chi_\ell + \chi_k + \varepsilon_{k,\ell} \tag{B.6}$$

where $\text{CPPMismatch}_{k,\ell}$ is defined using either method described in Section 5.1.2. We expect CPP mismatch with the frontier to inhibit biotechnology transfer; that is, we hypothesize that $\beta < 0$. Our estimates of Equation B.6 are displayed in Figure A7. Consistent with our main technology transfer results estimated at the country pair-by-crop level, we find that pathogen distance to frontier countries significantly inhibits biotechnology introduction in sub-Saharan Africa. While these estimates are necessarily less precise, given the smaller sample size and absence of data on the origin country, they tell a very similar story to our main analysis.

### B.5 Growth of US Biotechnology

Since the 1990s, the US agricultural biotechnology sector has produced a growing share of global innovation, likely driven by the advent and increased use of genetic modification. Figure A8 displays the relative growth of US patenting since 1990; the same trend for the EU is also reported, and does not show nearly as prominent an increase.

We investigate whether this shift in the geography of research affected the global distribution of production by disproportionately favoring producers in places where US technology—as opposed to European technology—was appropriate. For each country-crop pair, we measure the change in production (or area harvested) between the decade of the 1990s and the decade of the 2010s, and estimate:

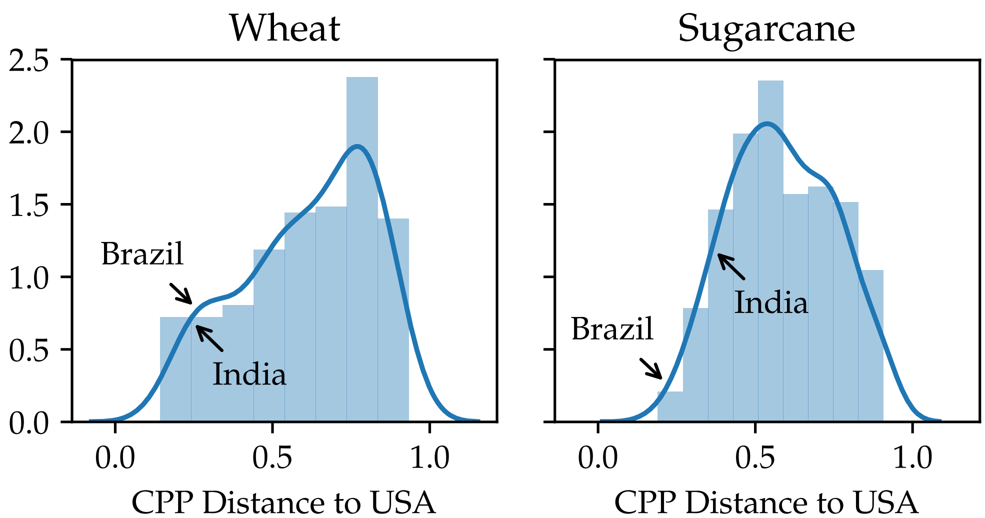
$$\Delta \log y_{k,\ell}^{90-10} = \beta_1 \cdot \text{CPP Mismatch}_{k,\ell}^{US} + \beta_2 \cdot \text{CPP Mismatch}_{k,\ell}^{EU} + \gamma \cdot \log y_{k,\ell}^{1990} + \chi_\ell + \chi_k + \varepsilon_{k,\ell} \tag{B.7}$$

Our first hypothesis is that $\beta_1 < 0$, capturing the effect of the rise of the US on production in places where US technology is more or less appropriate. Our second hypothesis is that $\beta_1 < \beta_2$, capturing the fact that since 1990, US technology has grown substantially more than European technology, so we would expect CPP mismatch with the US to be a more important determinant of productivity changes than CPP mismatch with Europe.

Estimates of (B.7) are reported in Table A17, and across specifications we find evidence of both hypotheses. $\beta_1 < 0$ and $\beta_2$ is close to zero and positive in all specifications. These estimates are less precise than our main results, and $\beta_1$ is statistically distinguishable from $\beta_2$ in just one of the four specifications. Nevertheless, dovetailing with Section 6.1, these findings suggest that global productivity differences are endogenous to the evolving landscape of technology development. As a result, geography does not have a fixed impact on development, but changing effects that can be shaped by the focus and direction of innovation.

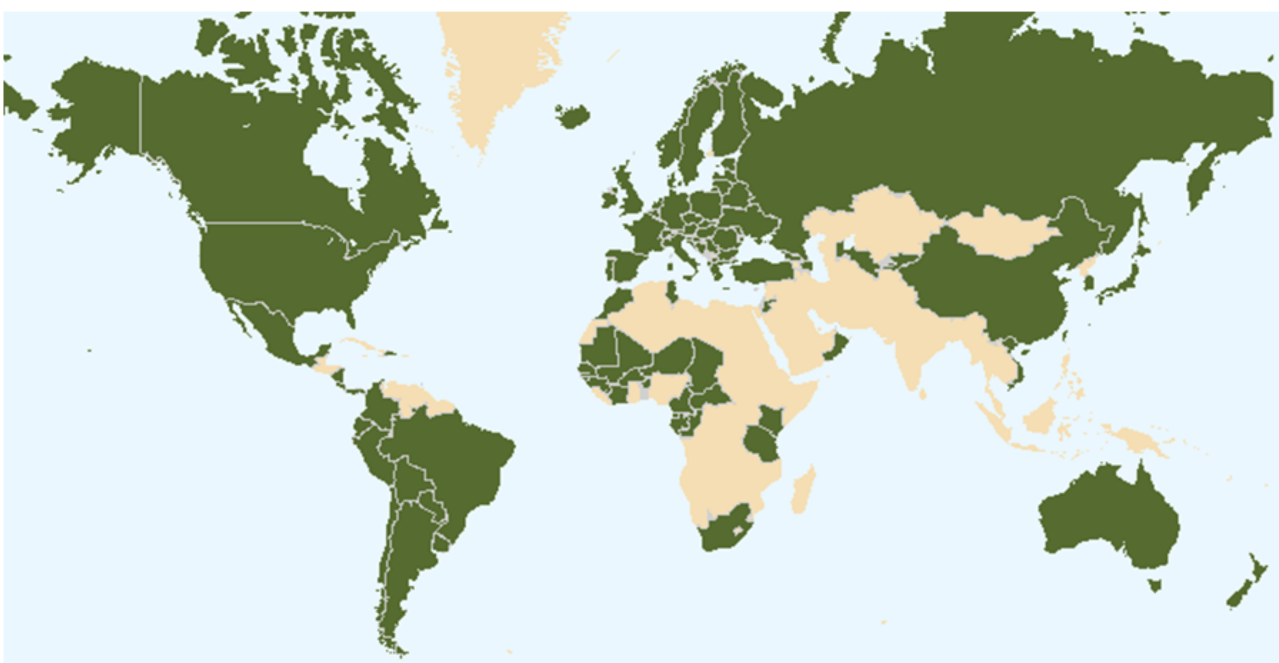
## C. Supplemental Figures and Tables

Figure A1: Example of CPP Mismatch Variation




*Notes*: Histogram (solid bars) and kernel density estimates (lines) for CPP Mismatch$_{\ell,\ell',k}$, where $\ell$ is the United States and $k$ is the crop indicated in each graph. Values for India and Brazil are labeled.

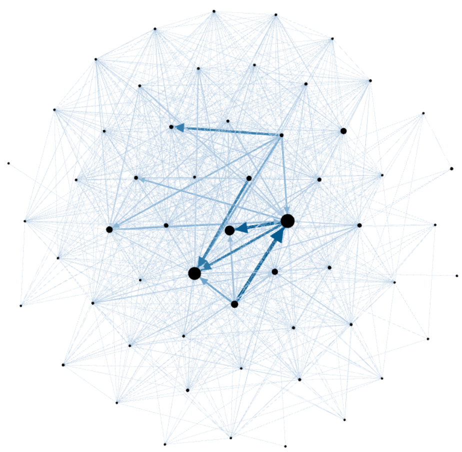
Figure A2: UPOV Compliant Countries



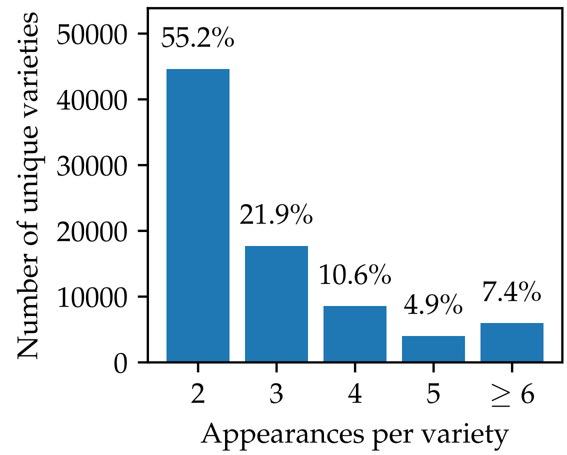
*Notes*: This figure denotes in green all UPOV member countries. This is the sample of countries for which we have data on biotechnology development and transfer.

Figure A3: Visualizing Variety Transfer

(a) Variety Transfer as a Directed Network



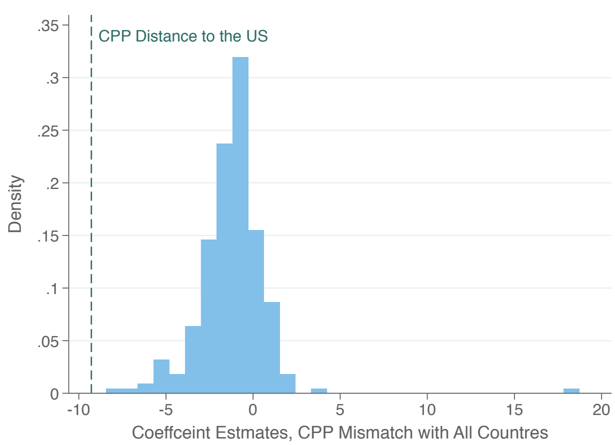
(b) Frequency of Occurrence for Varieties



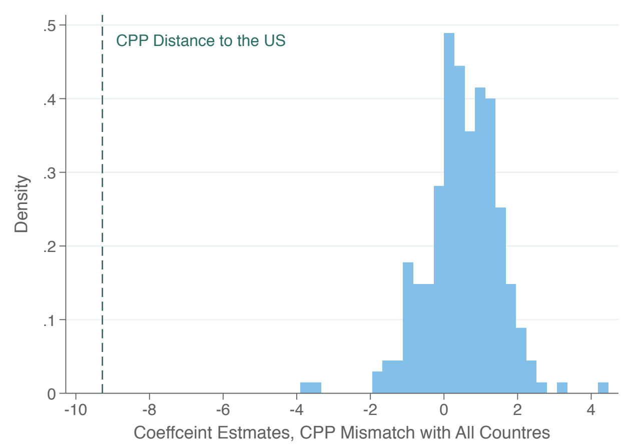

*Notes*: In (a), each node is a country sized in proportion to its total variety production and each edge is sized and colored in proportion to the number of varieties transferred. In (b), the percentages are in terms of unique varieties.

Figure A4: Falsification Test: CPP Mismatch with All Countries

(a) Unconditional



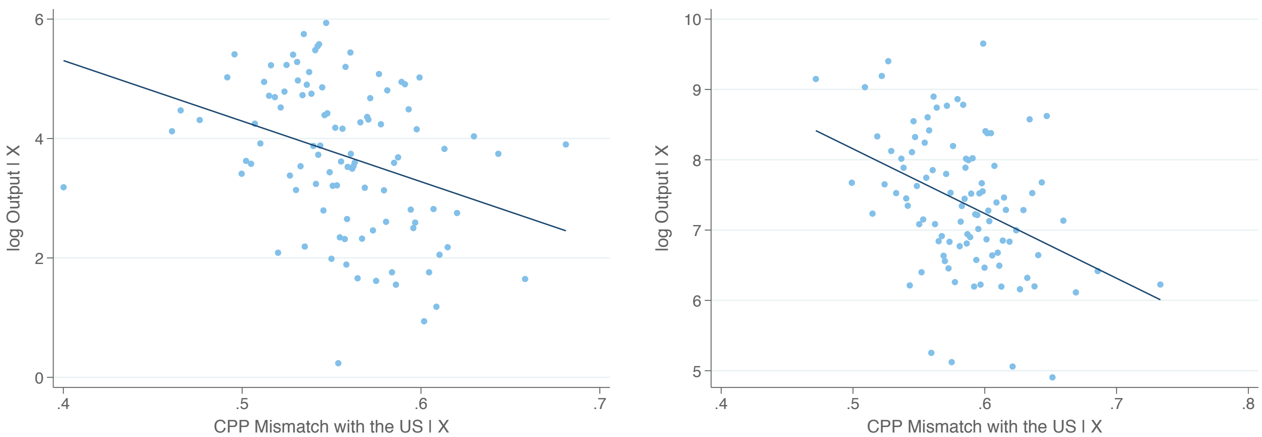

(b) Conditional on CPP Distance to the US




*Notes*: This figure displays histograms of the coefficient estimates of the relationship between CPP distance to each country separately and log of crop-level output. In A4a, CPP distance to each country is included on the right hand side of the regression alone (along with crop and country fixed effects) and A4b, CPP distance to the US is also included in the regression.

Figure A5: CPP Mismatch and Agricultural Output: Brazil and India Separately

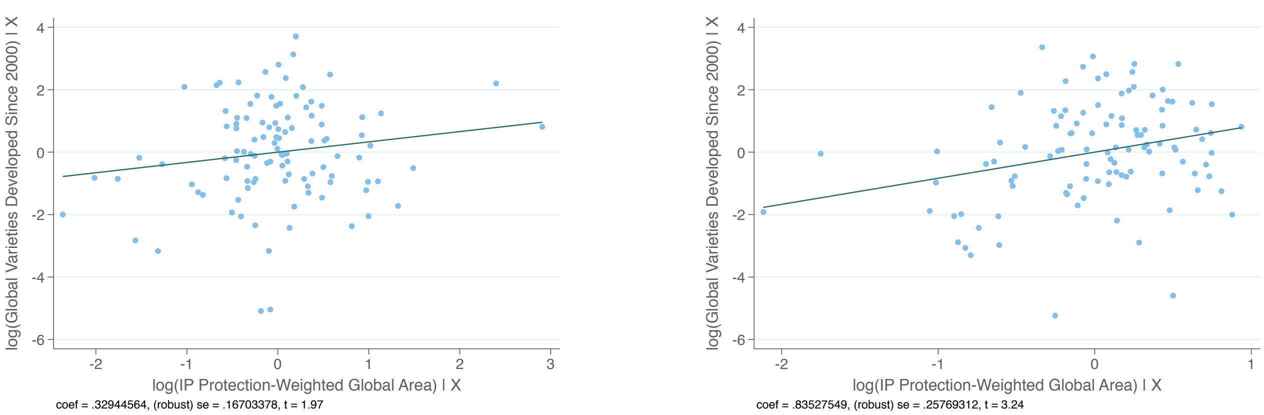
(a) India ($N = 384$) (b) Brazil ($N = 1,052$)




*Notes*: This figure displays binned partial correlation plots, after absorbing crop and state fixed effects, of our estimates of Equation (5.3), separately for India (A5a), where we estimate $\beta = -9.20$ (2.70), and Brazil (A5b), where we estimate $\beta = -10.15$ (5.17).

Figure A6: Bias in Global BioTech Development

(a) IP-Weighted Area and BioTech Development (b) GDP-Weighted Area and BioTech Development

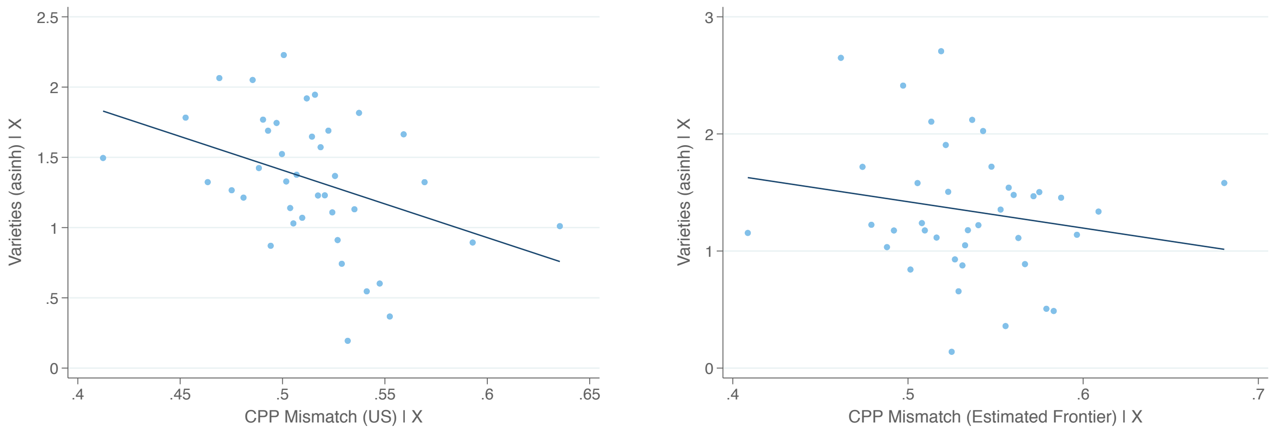



*Notes*: Partial correlation plots ($N = 107$) of our estimates of $\delta_2$ and $\delta_3$ from Equation (B.3). Both are estimated from the same regression, which also included a control for log of global planted area.

Figure A7: Pathogen Distance and Biotechnology Transfer to sub-Saharan Africa

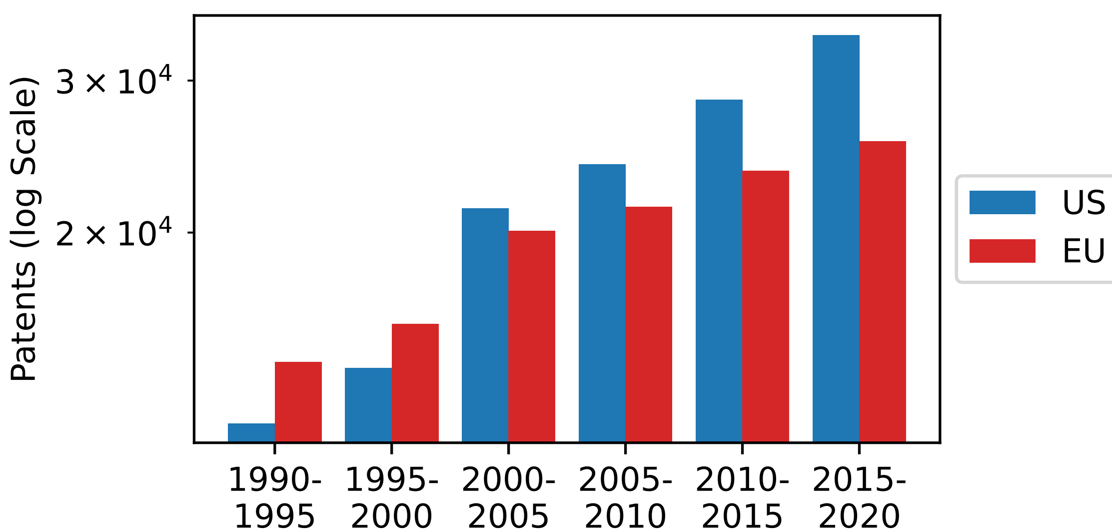
(a) CPP Mismatch with the US

(b) CPP Mismatch with Estimated Frontier Countries




*Notes*: This figure displays binned partial correlation plots, after absorbing country and crop fixed effects, of our estimates of Equation (B.6), both using pathogen distance to the US (left) and pathogen distance to the estimated frontier set (right). The number of observations is 345 in both sub-figures and standard errors are clustered by country.

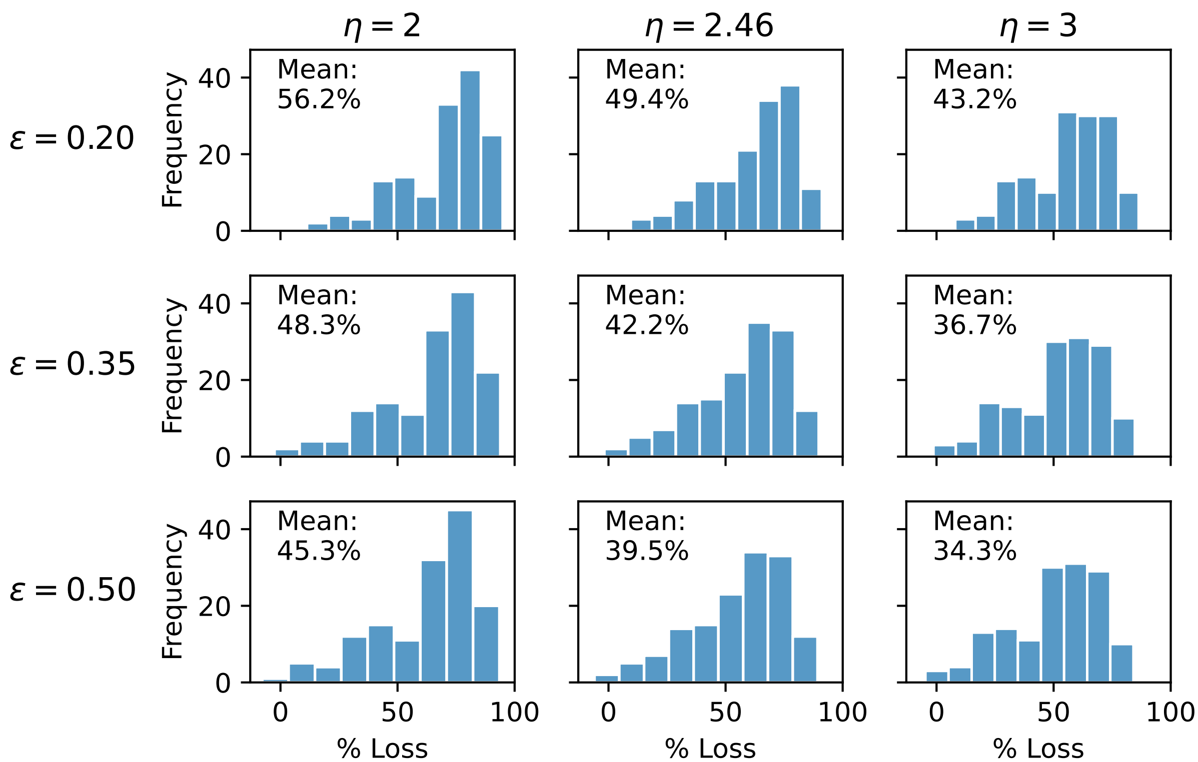
Figure A8: Growth in Agricultural Patented Technologies, Europe vs. the United States




*Notes*: Total number of patented agricultural technologies (i.e., in CPC class A01) in each five year period, comparing patents with assignees in the US to patents with assignees in the modern EU (as of 2018). Bars are the number of patents issued in the five year bin noted on the horizontal axis.

Figure A9: Sensitivity Analysis of Counterfactual Experiment

(a) Losses by Country




(b) Losses vs. Observed Productivity

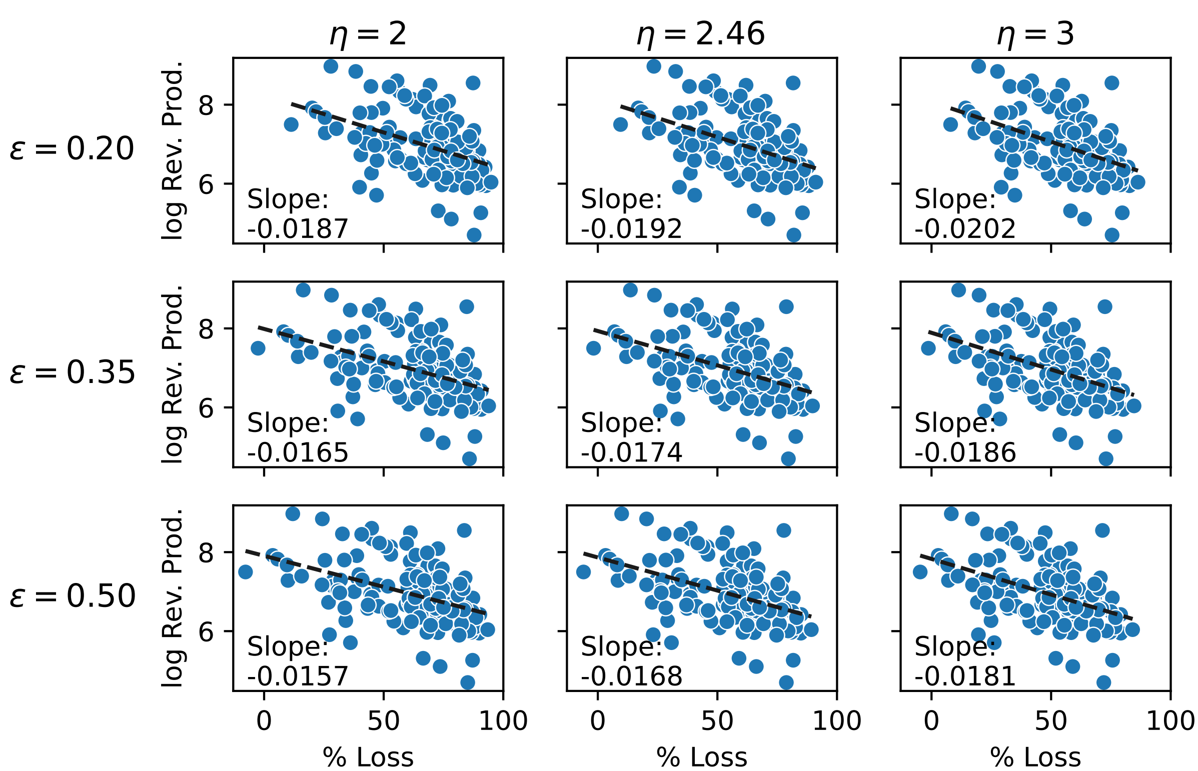
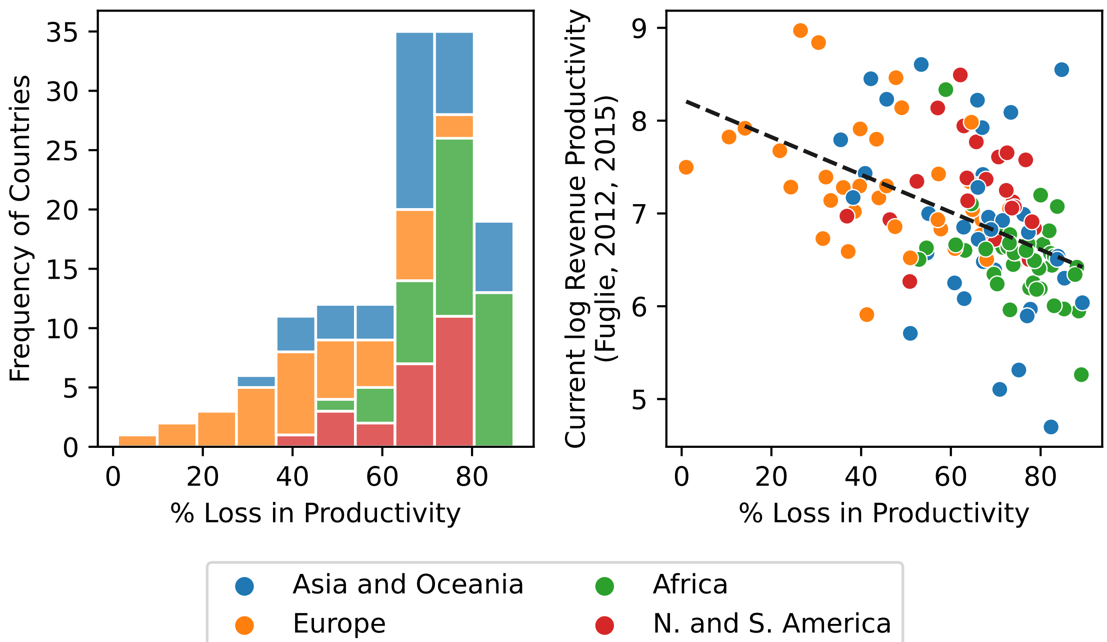


Figure A10: Causal Effects of Inappropriateness: CPP and Agro-Climatic Mismatch




*Notes*: This figure recreates Figure 5 under an experiment that removes inappropriate due to both CPP mismatch and Agro-Climatic mismatch. The left graph is a histogram of productivity losses from inappropriateness. The right graph is a scatterplot of productivity losses against observed productivity. The dashed line is a best-fit linear regression across countries. In each plot, colors indicate continents.

Table A1: Patenting Activity Directed Toward Local CPPs

| | (1) | (2) | (3) |
|---|---|---|---|
| | CPP-Specific Patents (asinh) | Any CPP-Specific Patent (0/1) | log CPP-Specific Patents |
| Local CPP | 0.0972*** | 0.0479*** | 0.181*** |
| | (0.0288) | (0.0106) | (0.0635) |
| Country Fixed Effects | Yes | Yes | Yes |
| CPP Fixed Effects | Yes | Yes | Yes |
| Observations | 492,422 | 492,422 | 8,557 |
| R-squared | 0.211 | 0.202 | 0.557 |

*Notes:* The unit of observation is a CPP-by-country pair. The dependent variable is the number of patents registered to inventors in the country and with the CPP's scientific name in the title, abstract, or patent description. Standard errors, clustered by country and CPP, are included in parentheses and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A2: Patenting Activity Directed Toward Local CPPs: Larger Effects in Rich Countries

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | CPP-Specific Patents (asinh) | Any CPP-Specific Patent (0/1) | log CPP-Specific Patents | CPP-Specific Patents (asinh) | Any CPP-Specific Patent (0/1) | log CPP-Specific Patents |
| Local CPP | 0.0720*** | 0.0395*** | 0.142* | 0.147*** | 0.0679*** | 0.172*** |
| | (0.0242) | (0.00887) | (0.0711) | (0.0418) | (0.0138) | (0.0521) |
| Local CPP x United States (0/1) | 1.002*** | 0.334*** | 0.394*** | | | |
| | (0.0274) | (0.0108) | (0.0825) | | | |
| Local CPP x log per-capita GDP (pre-period) | | | | 0.0860*** | 0.0366*** | 0.0492 |
| | | | | (0.0294) | (0.0101) | (0.0593) |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| CPP Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 492,422 | 492,422 | 8,557 | 364,144 | 364,144 | 8,478 |
| R-squared | 0.233 | 0.214 | 0.559 | 0.240 | 0.228 | 0.557 |

*Notes:* The unit of observation is a CPP-by-country pair. The dependent variable is the number of patents registered to inventors in the country and with the CPP's scientific name in the title, abstract, or patent description. GDP is computed at the country level from 1990-2000 and normalized by the global mean. Standard errors, clustered by country and CPP, are included in parentheses and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A3: CPP Mismatch Inhibits International Technology Transfer: Sensitivity Analysis

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | *Panel A: Dependent Variable is (asinh) Biotechnology Transfers* | | | | | | |
| CPP Mismatch (0-1) | -0.0624** | -0.113** | -0.0848*** | -0.0528** | -0.0572** | -0.0385** | -0.0443*** |
| | (0.0235) | (0.0467) | (0.0258) | (0.0227) | (0.0220) | (0.0186) | (0.0161) |
| | *Panel B: Dependent Variable is Any Biotechnology Transfer (0/1)* | | | | | | |
| CPP Mismatch (0-1) | -0.0275** | -0.0570** | -0.0373*** | -0.0226** | -0.0289*** | -0.0204** | -0.0239*** |
| | (0.0106) | (0.0218) | (0.0119) | (0.00998) | (0.0108) | (0.00855) | (0.00821) |
| | *Panel C: Dependent Variable is log Biotechnology Transfers* | | | | | | |
| CPP Mismatch (0-1) | -1.202*** | -0.937* | -0.935** | -1.198*** | -1.247*** | -1.888*** | -1.955*** |
| | (0.386) | (0.523) | (0.363) | (0.390) | (0.444) | (0.502) | (0.666) |
| Jaccard (1900, 1901) Distance Metric | | ✓ | | | | | |
| Broad CPP Presence Classification | | | ✓ | | | | |
| Control for bilaterial crop-level trade | | | | ✓ | | | |
| Control for log bilaterial distance x Crop FE | | | | | ✓ | | |
| Exclude country pairs <1000km apart | | | | | | ✓ | |
| Exclude country pairs <2000km apart | | | | | | | ✓ |
| Mean of CPP Distance Metric | 0.423 | 0.327 | 0.413 | 0.423 | 0.423 | 0.423 | 0.423 |
| Crop-by-Origin Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Crop-by-Destination Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Pair Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes:* The unit of observation is a crop-origin-destination. The dependent variable is noted in the header of each panel and the distiance metric, sample restrictiotn, and control set included in each specification is noted at the bottom of each column. Standard errors are double-clustered by origin and destination and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A4: CPP Mismach Reduces Area Harvested

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Dependent Variable is log Area Harvested | | | | | | | |
| | CPP Mismatch with the US | | | | CPP Mismatch with the Estimated Frontier | | | |
| CPP Mismatch (0-1) | -9.517*** | -12.08*** | -9.541*** | -7.855*** | -7.139*** | -7.020*** | -7.200*** | -5.837*** |
| | (1.212) | (2.892) | (0.595) | (0.635) | (0.941) | (0.725) | (0.437) | (0.496) |
| log(FAO-GAEZ-Predicted Output) | | 0.303*** | | | | 0.363*** | | |
| | | (0.0768) | | | | (0.0487) | | |
| *Included in LASSO Pool:* | | | | | | | | |
| Top CPP Fixed Effects | - | - | Yes | Yes | - | - | Yes | Yes |
| Ecological Features x Crop Fixed Effects | - | - | No | Yes | - | - | No | Yes |
| Controls in LASSO Pool | - | - | 335 | 3935 | | | 335 | 3935 |
| Crop Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6,675 | 2,268 | 6,683 | 5,908 | 6,469 | 2,268 | 6,474 | 5,748 |
| R-squared | 0.612 | 0.612 | | | 0.609 | 0.603 | | |

*Notes:* The unit of observation is a country-crop pair. Columns 1-4 use CPP mismatch with the US and columns 5-8 use CPP mismatch with the estimated set of technological leader countries. Columns 1-2 and 5-6 report OLS estimates and columns 3-4 and 7-8 report post double LASSO estimates. Country and crop fixed effects are included in all specifications, and included in the amelioration set in thet post-double LASSO specifications. The Top CPPs are defined as the top 200 CPPs defined by (i) the number of countries in whch they are present and (ii) the number of host crops that they infect. Since the two sets overlap, the total number is 335. The set of ecological features includes: temperature, precipitation, elevation, ruggedness, growng season days, soil acidity, soil clay content, soil silt content, soil coarse fragment volume, and soil water capacity. Standard errors are double-clustered by crop and state and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A5: CPP Mismach Reduces Exports and Increases Price Volatility

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | *Baseline Measure* | *Trade* | | *Producer Price Volatility* | | | |
| Dependent Variable: | log Output | log Exports | log Imports | Price SD (Norm. by Global Mean) | | log Price SD | |
| | *Panel A: CPP Mismatch wih the US* | | | | | | |
| CPP Mismatch (0-1) | -9.285*** | -8.768*** | 1.269 | 0.523*** | 0.317*** | 1.026*** | 0.671*** |
| | (1.199) | (1.200) | (1.295) | (0.126) | (0.109) | (0.237) | (0.224) |
| Observations | 6,926 | 5,495 | 5,854 | 4,580 | 4,559 | 4,580 | 4,559 |
| R-squared | 0.599 | 0.531 | 0.647 | 0.244 | 0.263 | 0.661 | 0.667 |
| | *Panel B: CPP Mismatch wih the Estimated Frontier Set* | | | | | | |
| CPP Mismatch (0-1) | -7.136*** | -5.386*** | -0.415 | 0.364*** | 0.212** | 0.628*** | 0.349** |
| | (0.959) | (0.877) | (0.871) | (0.101) | (0.0978) | (0.177) | (0.176) |
| Observations | 6,704 | 5,332 | 5,687 | 4,481 | 4,461 | 4,481 | 4,461 |
| R-squared | 0.600 | 0.535 | 0.649 | 0.243 | 0.262 | 0.662 | 0.668 |
| Crop Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Control for log Output | No | No | No | No | Yes | No | Yes |

*Notes:* The unit of observation is a crop-country pair. The dependent variable is listed at the top of each column and control set listed at the bottom. Standard errors are double-clustered by crop and country and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A6: CPP Mismach Reduces Output: Crop $\times$ Continent Fixed Effects

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Dependent Variable is log Output | | | | | | | |
| | CPP Mismatch with the US | | | | CPP Mismatch with the Estimated Frontier | | | |
| CPP Mismatch (0-1) | -8.809*** | -9.831*** | -8.780*** | -8.198*** | -8.780*** | -8.198*** | -6.999*** | -6.385*** |
| | (1.124) | (2.608) | (0.769) | (0.742) | (0.769) | (0.742) | (0.595) | (0.614) |
| log(FAO-GAEZ-Predicted Output) | | 0.239*** | | | | 0.273*** | | |
| | | (0.0704) | | | | (0.0770) | | |
| *Included in LASSO Pool:* | | | | | | | | |
| Top CPP Fixed Effects | - | - | Yes | Yes | - | - | Yes | Yes |
| Ecological Features x Crop Fixed Effects | - | - | No | Yes | - | - | No | Yes |
| Crop x Continent Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6,844 | 2,334 | 6,920 | 6,069 | 6,631 | 2,334 | 6,696 | 5,903 |
| R-squared | 0.680 | 0.694 | | | 0.679 | 0.689 | | |

*Notes:* The unit of observation is a country-crop pair. Columns 1-4 use CPP mismatch with the US and columns 5-8 use CPP mismatch with the estimated set of technological leader countries. Columns 1-2 and 5-6 report OLS estimates and columns 3-4 and 7-8 report post double LASSO estimates. Country and crop-by-continent fixed effects are included in all specifications, and included in the amelioration set in thet post-double LASSO specifications. The Top CPPs are defined as the top 200 CPPs defined by (i) the number of countries in whch they are present and (ii) the number of host crops that they infect. Since the two sets overlap, the total number is 335. The set of ecological features includes: temperature, precipitation, elevation, ruggedness, growng season days, soil acidity, soil clay content, soil silt content, soil coarse fragment volume, and soil water capacity. Standard errors are double-clustered by crop and country and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A7: CPP Mismach and Output: Additional Controls

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Dependent Variable is log Output | | | | | | | |
| | *Panel A: CPP Distance to the US* | | | | | | | |
| CPP Distance (0-1) | -9.122*** | -8.849*** | -9.573*** | -9.323*** | -9.186*** | -9.661*** | -10.10*** | -10.83*** |
| | (1.152) | (1.105) | (1.217) | (1.345) | (1.221) | (1.316) | (1.295) | (2.115) |
| Observations | 6,915 | 6,678 | 6,433 | 4,949 | 6,719 | 6,032 | 3,729 | 2,946 |
| R-squared | 0.600 | 0.632 | 0.612 | 0.634 | 0.614 | 0.626 | 0.671 | 0.786 |
| | *Panel B: CPP Distance to Estimated Frontier Set* | | | | | | | |
| CPP Distance (0-1) | -6.963*** | -6.838*** | -7.351*** | -7.206*** | -6.895*** | -7.172*** | -7.337*** | -7.250*** |
| | (0.934) | (0.879) | (1.029) | (1.065) | (0.980) | (1.011) | (1.058) | (1.743) |
| Observations | 6,693 | 6,458 | 6,227 | 4,765 | 6,499 | 5,838 | 3,631 | 2,864 |
| R-squared | 0.600 | 0.632 | 0.611 | 0.633 | 0.613 | 0.623 | 0.669 | 0.781 |
| Crop Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| log Per Capita GDP x Crop FE | No | Yes | No | No | No | No | No | Yes |
| Trade Share (% GDP) x Crop FE | No | No | Yes | No | No | No | No | Yes |
| Gini Coefficient x Crop FE | No | No | No | Yes | No | No | No | Yes |
| Share Arable Land x Crop FE | No | No | No | No | Yes | No | No | Yes |
| log Agricultural Value Added x Crop FE | No | No | No | No | No | Yes | No | Yes |
| R&D Share (% GDP) x Crop FE | No | No | No | No | No | No | Yes | Yes |

*Notes:* The unit of observation is a crop-country pair. Panel A uses CPP dstance to the US and Panel B uses CPP dstance to the estimated set of technological leader countries. Controls included in each specification are noted at the bottom of the column. Standard errors are double-clustered by crop and country and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A8: CPP Mismach Effects and Innovation

| | (1) | (2) |
|---|---|---|
| Dependent Variable: | log(BioTech Developed) | IP Protection (0/1) |
| $\beta^{\ell}$ | -0.584*** | -0.134*** |
| | (0.159) | (0.0173) |
| Observations (Countries) | 59 | 242 |
| R-squared | 0.173 | 0.250 |

Notes: The unit of observation is a country. log(BioTech Developed) is the (log of the) number of unique varieties developed in the country from 2000-2018. IP Protection (0/1) is an indicator variable that equals one if a country had UPOV compliant IP protection for plant biotechnology by 2000. $\beta^{\ell}$ refers to the coefficient estimate of the relationship between CPP mismatch with country $\ell$ and output. Both regressions are weighted by the inverse of the standard error of the estiamte of $\beta^{\ell}$. Robust standard errors are reported and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A9: Historical Green Revolution Breeding Sites

| (1) | (2) |
|---|---|
| Green Revolution Breeding Sites | |
| Crop | Site Location |
| Wheat | Mexico (CIMMYT) |
| Maize | Mexico (CIMMYT) |
| Sorghum | India (ICRISAT) |
| Millet | India (ICRISAT) |
| Beans | Colombia (CIAT) |
| Potatoes | Peru (CIP) |
| Cassava | Colombia (CIAT) |
| Rice | Philippines (IRRI) |

*Notes:* Column 1 reports the crops included in our analysis of the Green Revolution and column 2 reports the main breeding site during the Green Revolution for each crop, along with the corresponding IARC.

Table A10: Inappropriateness and the Efficacy of the Green Revolution

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Pct. Modern Variety Adoption | | | Δ log Output | Δ log Area Harvested |
| CPP Mismatch with GR Breeding Centers | -26.62*** | -96.20*** | -27.69*** | -2.642** | -2.501*** |
| | (9.155) | (27.17) | (9.492) | (1.052) | (0.881) |
| Crop Fixed Effects | Yes | Yes | - | - | - |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Crop x Continent Fixed Effects | - | - | Yes | Yes | Yes |
| Only Rice, Wheat, and Maize | No | Yes | No | No | No |
| Observations | 594 | 104 | 591 | 543 | 543 |
| R-squared | 0.406 | 0.677 | 0.471 | 0.419 | 0.419 |

*Notes:* The unit of observation is a country-crop pair. CPP mismatch for each crop is estimated as the CPP distance to the crop-specific Green Revolution main breeding center. All columns include crop and country fixed effects, as well as the pre-period value of the dependent variable, and columns 3-5 also include crop by continent fixed effects. In columns 1-3, the dependent variable is the change in percent (0-100) land area devoted to modern varieties between 1960 and 1980, and in columns 4 and 5 the dependent variable is the change in log output and log area harvested respectively, between the 1960s and the 1980s. Standard errors are double-clustered by country and crop-continent and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A11: Inappropriateness and the Efficacy of the Green Revolution: Timing and Geography

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Dependent Variable is $\Delta$ log Output | | | | | | |
| Sample: | Baseline Sample | | | All Africa | All South America | All Asia | All Europe |
| Time period: | 1960s-1980s | 1980s-2000s | 1990s-2010s | 1960s-1980s | 1960s-1980s | 1960s-1980s | 1960s-1980s |
| CPP Mismatch with GR Breeding Centers | -2.642** | -0.339 | -0.544 | -1.307 | -5.758** | -1.990 | 0.668 |
| | (1.052) | (0.832) | (0.783) | (0.808) | (1.903) | (1.372) | (1.516) |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Crop x Continent Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 543 | 540 | 538 | 277 | 83 | 207 | 118 |
| R-squared | 0.419 | 0.485 | 0.451 | 0.343 | 0.606 | 0.456 | 0.542 |

*Notes:* The unit of observation is a country-crop pair. CPP mismatch for each crop is estimated as the CPP distance to the crop-specific Green Revolution main breeding center. All columns include country and crop-by-continent fixed effects, as well as the pre-perod value of the dependent variable. The dependent variable is the change in log of crop output. The regresson sample as well as time period over which the change in output is calculated is listed at the top of each column. Standard errors are double-clustered by country and crop-continent in columns 1-3 and by country ni columns 4-7, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A12: CPP Mismatch Without Invasive Species: Baseline Estimates

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Technology Trasnfer | | | Technology Adoption | Output |
| Dependent Variable: | asinh Biotech Transfer | Any Biotech Transfer | log Biotech Transfer | Improved Seed (=1) | log Output |
| CPP Mismatch Without Invasive Species | -0.0712*** | -0.0304*** | -0.5451 | | |
| | (0.0241) | (0.0096) | (0.34) | | |
| CPP Mismatch with the Frontier Without Invasive Species | | | | -0.248*** | -6.335*** |
| | | | | (0.0743) | (0.948) |
| Crop-by-Origin Fixed Effects | Yes | Yes | Yes | - | - |
| Crop-by-Destination Fixed Effects | Yes | Yes | Yes | - | - |
| Country Pair Fixed Effects | Yes | Yes | Yes | - | - |
| Country Fixed Effects | - | - | - | Yes | Yes |
| Crop Fixed Effects | - | - | - | Yes | Yes |
| Observations | 202,154 | 202,154 | 5,752 | 115,397 | 6,858 |
| R-squared | 0.4397 | 0.3831 | 0.7965 | 0.213 | 0.584 |

*Notes:* The unit of observation is a crop-origin-destination in columns 1-3 and a crop-country pair in columns 4-6. Standard errors are double-clustered by origin and destination in columns 1-3, clustered y crop-country in columns 4-5, and double clustered by crop and country in column 6. CPP mismatch with the frontier is computed as CPP mismatch with the US. In all cases, the independent variable is constructed after excluding invasive CPPs. The fixed effects included in each specification are noted at the bottom of each column. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A13: Agro-climatic Mismatch and Technology Transfer

| | (1) | (2) | (3) |
|---|---|---|---|
| | Dependent Variable is (asinh) Biotechnology Transfers | | |
| CPP Mismatch (0-1) | -0.0783** | -0.0737** | -0.0752** |
| | (0.0314) | (0.0309) | (0.0311) |
| *Mismatch in:* | | | |
| Temperature | | -0.0107* | |
| | | (0.00619) | |
| Precipitation | | -0.0141* | |
| | | (0.00807) | |
| Elevation | | -0.00589* | |
| | | (0.00311) | |
| Ruggedness | | -0.000652 | |
| | | (0.00246) | |
| Soil Clay Content | | -0.00596 | |
| | | (0.00568) | |
| Soil Silt Content | | 0.00342 | |
| | | (0.00575) | |
| Soil Coarse Fragment Content | | 0.000883 | |
| | | (0.00318) | |
| Soil pH | | -0.00825** | |
| | | (0.00355) | |
| Growing Season Length | | -0.00453 | |
| | | (0.00519) | |
| Available Water Capacity | | -0.00561 | |
| | | (0.00466) | |
| Overall Agro-Climatic Mismatch | | | -0.0412*** |
| | | | (0.0129) |
| p-value joint significance | - | 0.007 | - |
| Crop-by-Origin Fixed Effects | Yes | Yes | Yes |
| Crop-by-Destination Fixed Effects | Yes | Yes | Yes |
| Country Pair Fixed Effects | Yes | Yes | Yes |
| Observations | 153,038 | 153,026 | 153,038 |
| R-squared | 0.464 | 0.464 | 0.464 |

Notes: The unit of observation is a crop-origin-destination. Mismatch in agro-climatic features is estimated by first calcualting the value of each characteristic in the land area devoted to each crop in each country, as recorded by the EarthStat database. The agro-climatic index in column 3 is constructed as a sum of the normalized values of the characteristics listed in column 2. Standard errors are double-clustered by origin and destination and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A14: Agro-climatic Mismatch and Agricultural Output

| | (1) | (2) |
|---|---|---|
| | Dependent Variable is log Output | |
| CPP Mismatch (0-1) | -7.511*** | -6.682*** |
| | (1.361) | (1.344) |
| Overall Agro-Climatic Mismatch | | -1.222*** |
| | | (0.318) |
| Crop Fixed Effects | Yes | Yes |
| Country Fixed Effects | Yes | Yes |
| Observations | 4,881 | 4,881 |
| R-squared | 0.574 | 0.580 |

Notes: The unit of observation is a crop-country pair. Mismatch in agro-climatic features is estimated by first calcualting the value of each characteristic in the land area devoted to each crop in each country, as recorded by the EarthStat database. The agro-climatic index is constructed as a sum of the normalized values of the individual characteristics. Standard errors are double-clustered by crop and country and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A15: Correlation Matrix: All Ecological Mismatch Measures

| Difference in: | CPPs | Temp. | Precip. | Elevation | Rugged. | Soil Clay Content | Soil Silt Content | Coarse Frag. Content | Soil pH | Growing Season Length | Available Water Capacity |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CPPs | 1.0000 | | | | | | | | | | |
| Temp. | 0.2356 | 1.0000 | | | | | | | | | |
| Precip. | 0.1061 | 0.2121 | 1.0000 | | | | | | | | |
| Elevation | 0.1578 | 0.0104 | -0.0405 | 1.0000 | | | | | | | |
| Rugged. | 0.1726 | -0.0382 | 0.05 | 0.5052 | 1.0000 | | | | | | |
| Soil Clay Content | 0.0374 | 0.1602 | 0.146 | -0.0074 | -0.0096 | 1.0000 | | | | | |
| Soil Silt Content | 0.1807 | 0.3564 | 0.0236 | 0.0402 | -0.1209 | 0.0966 | 1.0000 | | | | |
| Soil Coarse Fragement Content | 0.1045 | 0.0697 | 0.0188 | 0.3407 | 0.5595 | -0.0999 | -0.1013 | 1.0000 | | | |
| Soil pH | 0.0793 | 0.0829 | 0.4994 | -0.0082 | 0.0128 | 0.1087 | 0.0326 | -0.0001 | 1.0000 | | |
| Growing Season Length | 0.084 | 0.1186 | 0.5092 | -0.0121 | 0.009 | 0.0216 | 0.0275 | 0.0001 | 0.4116 | 1.0000 | |
| Available Water Capacity | 0.1375 | 0.1829 | 0.099 | 0.0126 | -0.0466 | 0.3531 | 0.3893 | -0.0966 | 0.0906 | 0.0665 | 1.0000 |

*Notes:* This table presents a correlation matrix among all individual measures of ecological distance to the frontienr, including CPP distance to the frontier. The additional characteristics are: tempearture, precipitation, elevation, ruggedness, soil clay content, soil silt content, soil coarse fragement content, soil pH, growing season length, and available water capacity. Each cell reports a pairwise correlation coefficient.

Table A16: Global Bias of Technology Development: Crop-by-Country Estimates

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable: | asinh(BioTech Since 2000) | | | | | |
| asinh(Local Area) | 0.227*** | 0.213*** | 0.209*** | 0.204*** | 0.204*** | 0.155*** |
| | (0.0125) | (0.00986) | (0.0112) | (0.00977) | (0.00982) | (0.00842) |
| asinh(Global Area) | | 0.0565*** | -0.0451 | -0.0155 | -0.0551 | |
| | | (0.0208) | (0.0540) | (0.0310) | (0.0459) | |
| asinh(GDP-Weighted Area) | | | 0.0925 | | 0.0512 | |
| | | | (0.0606) | | (0.0620) | |
| asinh(IP-Weighted Area) | | | | 0.0814*** | 0.0625* | |
| | | | | (0.0309) | (0.0369) | |
| Country Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Crop Fixed Effects | No | No | No | No | No | Yes |
| Observations | 6,758 | 6,758 | 6,758 | 6,758 | 6,758 | 6,758 |
| R-squared | 0.495 | 0.501 | 0.505 | 0.506 | 0.507 | 0.600 |

*Notes:* The unit of observation is a crop-by-country pair. The dependent variable is the number of varieties developed in the country for the crop since 2000. Standard errors, clustered by crop, are included in parentheses and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A17: Growth of US Biotechnology and Global Production

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Δ log Output | | Δ log Area Harvested | |
| CPP Mismatch with the US | -0.999* | -0.974* | -1.004** | -1.044* |
| | (0.520) | (0.572) | (0.502) | (0.533) |
| CPP Mismatch with the EU | 0.644 | 0.251 | 0.352 | 0.222 |
| | (0.512) | (0.531) | (0.529) | (0.534) |
| Crop Fixed Effects | Yes | - | Yes | - |
| Country Fixed Effects | Yes | Yes | Yes | Yes |
| Crop x Continent Fixed Effects | - | Yes | - | Yes |
| *p-value,* Dist US - Dist EU | 0.097 | 0.249 | 0.172 | 0.216 |
| Observations | 6,414 | 6,338 | 6,183 | 6,107 |
| R-squared | 0.281 | 0.366 | 0.262 | 0.353 |

*Notes:* The unit of observation is a country-crop pair. Both CPP mismatch with the US and CPP mismatch with the EU are included in all specifications. All columns include crop and country fixed effects, as well as the pre-period value of the dependent variable, and columns 2 and 4 also include crop by continent fixed effects. In columns 1-2, the dependent variable is the change in log output from the 1990s to 2010s and in columns 3-4 it is the change in log area harvested from the 1990s to 2010s. Standard errors are double-clustered by country and crop and *, **, and *** indicate significance at the 10%, 5%, and 1% levels.