Optimal Gradualism*

Nils H. Lehr
Boston University

Pascual Restrepo
Boston University

October 23, 2022

Abstract

This paper studies how gradualism affects the welfare gains from trade, technology, and reforms. When people face adjustment frictions, gradual shocks create less adverse distributional effects in the short run. We show that there are welfare gains from inducing a more gradual transition via temporary taxes, and provide formulas for the gains from gradualism and optimal short-run taxes on trade and technology. Our formulas account for the possibility that reallocation effort is endogenous and responds to policy, and for the existence of public assistance programs. Using these formulas, we compute the optimal temporary taxes needed to mitigate the distributional consequences of rising import competition from China and the deployment of automation technologies substituting for routine jobs. Our formulas can also be used to compute the optimal timing of economic reforms or trade liberalizations, and we apply them to study Colombia's trade liberalization in 1990—a prominent example where optimal policy called for a more gradual reform.

*Restrepo thanks the National Science Foundation for its support under award No. 2049427. We thank David Autor and Marcela Eslava for sharing their data and providing feedback on this project.

Technological progress, trade, and economic reforms generate periods of adjustment during which some people fall behind, loose their jobs, experience wage declines, and see their livelihoods disrupted.[1] Even if technology and trade are positive developments in the long-run, dealing with short-run disruption costs for some households remains an important policy concern, especially in the wake of rapid changes in the economy.[2]

Existing evidence points to large disruption costs. Autor et al. (2014) document that an average worker in an industry with a high level of exposure to Chinese import competition experienced a cumulative income loss relative to unexposed workers equal to half their annual earnings in 1990 over the 1992–2007 period. Cortes (2016) shows that workers who in 1985 held routine jobs — those that can be more easily automated — experienced a subsequent decline in wages of 20% by 2007 relative to similar workers in other occupations.

How should policy respond during these periods of adjustment? Do short-run disruption costs imply that more gradual advances in technology and trade are preferable?

Our main point is that the short-run disruption costs brought by technological change, trade, and reforms are more severe when these changes take place abruptly, without giving workers opportunity to adjust. Abrupt change, thus, creates gains from gradualism and justifies temporary taxes on new technologies and trade or embracing gradual reforms.

Our main contribution is to provide formulas for the welfare gains from gradualism and the optimal path for taxes on new technologies and trade that capture these insights. We showcase these formulas in a calibrated version of our model that matches the empirical estimates of Autor et al. (2014) for trade and Cortes (2016) for the automation of routine jobs. The model calls for temporary taxes on trade and automation technologies of 10%–15% which are then phased out over time. We also use our formulas to study Colombia's trade liberalization in 1990 and show that optimal policy called for a more gradual reform.

We derive these formulas in a general model of worker displacement by trade or technology. Ex-ante identical workers are allocated across islands a-la Lucas and Prescott (1974) and each island represents segments of industries being disrupted by international trade

[1] For evidence in the context of trade, see Autor and Dorn (2013); Autor et al. (2014). For evidence in the context of automation technologies, see Cortes (2016); Adão et al. (2021); Acemoglu and Restrepo (2020, forthcoming). Finally, see Goldberg and Pavcnik (2005) for evidence on how dismantling trade protection reduces the relative wages of workers in exposed industries.

[2] In the US, industrial robots installations and imports from China tripled in a few years (see Autor et al., 2013; Acemoglu and Restrepo, 2020, respectively), and the share of e-commerce in retail went from 0.6% to 10% from 1999 to 2019 (see US Census, 2022). As Erik Brynjolfsson and Andrew McAffe put it in *The Second Machine Age*, "People are falling behind because technology is advancing so fast and our skills and organizations aren't keeping up" (Brynjolfsson and McAfee, 2014). Managing short-run disruptions is also a key policy concern when it comes to policy reforms (see, for example, Rodrik, 1995).

(e.g., low-cost apparel or household electronics) or jobs being automated by new technologies (e.g., welding or data-entry clerks). At time t_0, a new technology arrives, capable of replacing workers in disrupted islands by producing the same output at lower costs. These costs further decline over time as the technology improves exogenously, capturing advances in the production of automation equipment and software capable of substituting for workers at some tasks (as in Acemoglu and Restrepo, forthcoming) or improvements in Chinese productivity leading to the disruption of some segments of US industries (as in Caliendo et al., 2019; Galle et al., 2022). Technology causes real wages at disrupted islands to fall over time and real wages at other islands to increase. As in Alvarez and Shimer (2011), workers reallocate at a constant rate $\alpha > 0$, such that the transition features a temporary decline in the real wage and consumption of some workers and higher real wages for all in the long run.[3] We interpret the slow adjustment process as reflecting the time it takes for workers to reallocate, find new jobs, or acquire the skills required in other jobs.

Using this model, we provide analytical answers to the two questions raised above:

Given a path for technological progress, should the government induce a more gradual adjustment via temporary taxes on new technologies?

Using a variational approach, we derive formulas for the optimal tax path on new technologies and show that the optimum involves a temporary increase in taxes which is then fully phased-out over time. Temporary optimal taxes are positive even if technology, trade, and reforms make everyone better off in the long run. Taxes on new technologies should be higher when people in disrupted islands experience a large drop in consumption during the transition, which is itself linked to the decline in income documented by Autor et al. (2014) for the China Shock and Cortes (2016) for the automation of routine jobs.

Our formulas account for the possibility that taxing new technologies might generate adverse incentives for reallocation, which creates a motive for a faster phase out of the initial tax. Our formulas also extend to an scenario in which the planner can set temporary assistance programs that transfer income to households experiencing income losses at some replacement rate. We show that taxes on new technologies are justified if assistance programs can only be conditioned on income but not on work effort or islands.[4] The logic is the

[3] Our model is designed to capture the effects of labor-replacing technologies or technologies that work by substituting workers at some of their existing roles. We see trade and automation technologies as working in this way. These technologies have the potential to reduce wages of displaced workers and raise wages of all other workers. Other developments, such as factor-neutral improvements in technology, or technologies that directly complement skilled workers do not fit into our description.

[4] If island-specific lump-sum transfers or island-specific wage subsidies were available, redistribution can be done without distorting production. This is a direct implication of the Second Welfare Theorem or

same as in Naito (1999) and Costinot and Werning (forthcoming): taxing new technologies has tagging value because it assists workers affected by these exogenous disruptions and not those who reduced their work effort to take advantage of assistance programs.

Conditional on government policy, does the economy benefit from more rapid technological advances along the transition?

Here too, we use a variational approach to compute the welfare gains from moving to a counterfactual world where advances in technology are more gradual. We refer to these as the gains (or costs) from *technological gradualism*. The gains from technological gradualism are positive when technology advances rapidly and displaced workers experience large drops in consumption. However, with optimal taxes in place, there are no gains from technological gradualism and faster technological progress is always welcomed.

Applications: We apply our framework to study the automation of routine jobs, the China Shock, and Colombia's 1990 trade liberalization. We adopt an utilitarian welfare function and calibrate the model to match the evidence in Cortes (2016) on the automation of routine jobs and in Autor et al. (2014) on the China Shock. The evidence points to limited opportunities for reallocation and implies low values for the reallocation rate α of 2.5% per year for routine jobs and 1.8% per year for the China Shock. In addition, we back out the underlying path for technology from data on occupational wages or import shares.

We find that the automation of routine jobs had a sizable negative welfare effect of 6%–8% on workers in disrupted islands, depending on assumptions about savings and initial assets. This is driven by a short-run income decline of 12% from 1985 to 2000, which recovers by 2025. These short-run disruption costs justify an optimal tax on automation technologies of the order of 10%–12.5% over 1985–1995, which is then phased out slowly and reaches a level of 4% by 2020 in the least gradual scenario.

The China Shock had a sizable negative welfare effect of 15%–19% on workers in disrupted islands, depending on assumptions about savings and initial assets, though these islands account for only 1.6% of the US workforce. Welfare losses are driven by a short-run income decline of 35% from 1991 to 2004. These short-run disruption costs justify an optimal tax on Chinese imports of the order of 10%–15% over 1991–2000, which is then phased out slowly and reaches a level of 3% by 2020 in the least gradual scenario.

the taxation principles in Diamond and Mirrlees (1971). In practice, these additional (and more desirable policy tools) might be limited, since identifying workers whose livelihoods were disrupted by technology and trade as opposed to regular economic churn and sectoral fluctuations might be challenging.

In both applications, we find no gains from more technological gradualism, even absent taxes on technologies or trade. The model suggests that while it is optimal to tax these technological developments in the short run, society would not benefit from moving to a counterfactual world where trade and automation advance more slowly for exogenous technological reasons. These statements are not in contradiction: taxing technologies generates a more gradual path for wages and additional tax revenue; while a more gradual technological path generates no tax revenue.

In a final application, we use our formulas to compute the optimal trade liberalization policy for Colombia. In 1990, Colombia embarked in a rapid and ambitious program of trade liberalization, reducing effective tariffs by 37.5% in two years. We calibrate our model to match the immediate increase in imports following the reform and the drop in wages in previously protected industries estimated by Goldberg and Pavcnik (2005), which points to small values of α of 3% per year. Optimal policy calls for a more gradual reform, with tariffs remaining at a fourth of their initial level by 2000—10 years after the reform started. Reallocation rates of 20% per year—one order of magnitude higher than what we estimate—are needed to justify Colombia's swift drop in tariffs.

Related literature: Our optimal tax formulas in Propositions 3 and 4 relate to the tariff formula in Grossman and Helpman (1994) and the formula for optimal taxes on new technologies in Propositions 1 and 3 of Costinot and Werning (forthcoming). Grossman and Helpman (1994) and Helpman (1997) focus on redistribution via tariffs across workers specialized in different industries. Our optimal tax formula shares the same structure as theirs and generalizes it to a dynamic environment where workers reallocate over time.[5]

We borrowed the idea of characterizing optimal taxes using variational arguments from Costinot and Werning (forthcoming), and extended their arguments to a dynamic setting.[6] Despite methodological similarities, the problem solved by Costinot and Werning differs from ours. They are interested in how taxing a new technology can help reduce underlying inequalities between ex-ante different people. Technologies that permanently reduce wages at the bottom of the income distribution relative to the top have "tagging" value. Taxing these technologies achieves a better distribution of income than using income taxes alone, an insight that goes back to Naito (1999). The problem we study is different and complementary. In our model, the purpose of taxing new technologies is easing the transition for

[5]The formulas also differ in that Grossman and Helpman assume a quasi-linear aggregator across islands (and their weights emerge from lobbying and not necessarily from welfare considerations).

[6]Variational arguments have been used extensively to characterize properties of optimal income tax schedules (Saez, 2001; Tsyvinski and Werquin, 2017; Saez and Stantcheva, 2016).

workers left behind in a world where winners and losers are ex-ante identical. This is why, in our formulas, optimal taxes are linked to the short-run decline in income for exposed workers relative to similar non-exposed workers (i.e., the incidence regressions in Autor et al., 2014; Cortes, 2016, used here), and do not depend on how trade or robots affect incomes at different points of the income distribution (i.e., the quantile regressions in Acemoglu and Restrepo, 2020; Chetverikov et al., 2016, used by Costinot and Werning). We interpret the formulas in Costinot and Werning (forthcoming) as prescribing long-run taxes designed to improve the distribution of income, and our formulas as prescribing short-run taxes designed to ease disruption costs.

Our paper also contributes to a recent literature on the optimal taxation of automation motivated by distributional considerations (Thuemmel, 2018; Guerreiro et al., 2021; Donald, 2022) or inefficiencies (Acemoglu et al., 2020; Beraja and Zorzi, 2022). Thuemmel (2018); Guerreiro et al. (2021) show that non-zero taxes on robots are justified even when income taxes are available as an additional tool for redistribution, in line with Naito (1999).[7] Like us, Guerreiro et al. (2021) emphasize that optimal taxes on robots are positive along the transition and zero in the long run, when affected cohorts of workers retire from the labor market. Beraja and Zorzi (2022) also argue for temporary taxes on automation technologies, though in their case taxation is motivated by efficiency considerations: firms do not internalize the fact that their decision to automate push displaced workers against their borrowing constraint, which generates excessive automation.[8] We contribute to this literature by providing intuitive and general formulas for optimal taxes on automation technologies that provide a tight link between the theory and the empirical evidence and identify the key features of the data that inform optimal taxes. Finally, we show that the same formulas can be applied to understanding how trade competition should be handled during a transition period and how economic reforms should be conducted.

Finally, we contribute to the literature on the optimal timing of reforms and trade liberalization, going back to Mussa (1984) and with subsequent contributions by Edwards and van Wijnbergen (1989); Karp and Paul (1994); Rodrik (1995); Bond and Park (2002); Chisik (2003). In his seminal work, Mussa argued that "a general case for gradualism in trade liberalization can be based on a desire to limit the income and wealth losses sustained by owners of resources initially employed in protected industries," which is the rationale

[7]A complementary line of work studies the compensation of displaced workers via changes in income taxes (see Antràs et al., 2017; Tsyvinski and Werquin, 2017) and derives formulas for welfare that account for distributional considerations when these optimal taxes are in place.

[8]In their extensions, Beraja and Zorzi consider the role of redistribution and show that this leads to higher taxes on new technologies, which is in line with our results.

for gradualism studied in this paper.

Roadmap: Section 1 introduces our model and characterizes the transitional dynamics following advances in trade or labor-replacing technologies. Section 2 derives formulas for optimal taxes and the gains from technological gradualism. Sections 3, 4, and 5 apply our framework to the automation of routine jobs in the US, the China Shock, and Colombia's trade liberalization, respectively. Proofs and derivations are in the Appendix.

1 A Model of Economic Disruptions

Status quo: Consider an economy with a mass 1 of workers and a set of islands $x \in \mathcal{X}$. Each island produces a good $y_{x,t}$, which combines with the output of other islands into a final numeraire good y_t according to a constant returns to scale production function

$$y_t = f\left(\{y_{x,t}\}_{x \in \mathcal{X}}\right).$$

Initially, islands produce these goods with labor, so that $y_{x,t} = \ell_{x,t}$, where $\ell_{x,t}$ denotes the mass of workers in island x. We assume that the initial allocation of workers $\ell_{x,0}$ across islands before the shock equated all island wages to a common level $\bar{w}$. This can be thought of as the steady state of the reallocation process introduced below.

The disruption: At time $t = 0$ a new technology arrives. For a subset of *disrupted islands* $x \in \mathcal{D} \subset \mathcal{X}$, it becomes possible to produce their goods using a new technology embodied in capital $k_{x,t}$. New capital can be produced from $1/A_{x,t}$ units of the final good, where $A_{x,t}$ is the productivity of the new technology, which increases over time and converges to A_x.

Following the arrival of new technologies, the production of $y_{x,t}$ becomes

$$y_{x,t} = \begin{cases} \ell_{x,t} + k_{x,t} & \text{if } x \in \mathcal{D} \\ \ell_{x,t} & \text{if } x \notin \mathcal{D}. \end{cases}$$

The disruption leads to permanent changes in island wages $w_{x,t}$, which prompt workers to reallocate with Poisson probability $\alpha_x > 0$ to an island of their choice.

Taxes: The government sets taxes $\tau_{x,t}$ on new technologies, raising revenue

$$T_t = \sum_{x \in \mathcal{D}} \tau_{x,t} \cdot \frac{k_{x,t}}{A_{x,t}},$$

which gets redistributed in a lump sum way. We first consider a baseline version of our model where the government has no other tools for redistribution or assistance, and study these tools in our extensions.

Households: Households in island x are identical before the shock and hold assets $a_{x,0}$. After the shock, households at island x make consumption decisions to maximize

$$U_{x,0} = \mathbb{E}\left[\int_0^\infty e^{-\rho t} \cdot u(c_{x,t}) \cdot dt\right] - \kappa(\alpha_x)$$

subject to some budget constraints that we left unspecified, but that could capture various scenarios, ranging from hand-to-mouth, to perfect risk sharing within islands.

For our purposes, it suffices to work with households indirect utility function

$$U_{x,0} = \mathcal{U}\left(\{w_{x',t} + T_t\}_{x' \in \mathcal{X}}, a_{x,0}; \alpha_x\right) - \kappa_x(\alpha_x),$$

which gives the maximum expected utility for a household in island x facing a path of incomes $\{w_{x',t} + T_t\}_{t=0}^{\infty}$ across islands and that reallocates at a rate α_x.[9]

The term $\kappa_x(\alpha_x)$ captures reallocation costs. We consider two cases. With *exogenous reallocation,* $\alpha_x > 0$ is fixed and we set $\kappa_x(\alpha_x) = 0$. With *endogenous reallocation,* α_x is chosen by households to maximize $\mathcal{U}\left(\{w_{x',t} + T_t\}_{x' \in \mathcal{X}}, a_{x,0}; \alpha_x\right) - \kappa_x(\alpha_x)$.

1.1 Transitional Dynamics and Equilibrium

We impose two assumptions on f, which are satisfied by commonly used aggregators.

Assumption 1 (Symmetry among undisrupted islands) *For all islands $x', x'' \notin \mathcal{D}$ and any island $x \in \mathcal{D}$, we have*

$$\frac{\partial^2 f}{\partial y_x \partial y_{x'}} = \frac{\partial^2 f}{\partial y_x \partial y_{x''}}.$$

This assumption ensures that new technologies benefit all undisrupted islands equally. It implies a common wage w_t paid at undisrupted islands along the transition. This allows us to study redistribution between undisrupted and disrupted islands—the winners and

[9]Our formulation assumes that either we are in a small open economy and the interest rate is fixed, or households are hand to mouth and do not save nor borrow. This is why the interest rate does not affect indirect utilities.

losers of trade, technological progress, or reforms—and abstract from redistribution between winners (i.e., software engineers benefiting more than economists from the use of software to automate sales jobs).

Let $c^f(p)$ denote the unit cost function associated with the aggregator f. With some abuse of notation, we denote by $c^f(\{w_x\}_{x\in\mathcal{D}}, w)$ the price of the final good when the price of the island x output is w_x for $x \in \mathcal{D}$ and w for other islands. Also, we denote by c^f_x and c^f_w the derivatives of this cost function with respect to w_x and w respectively. Formally:

$$c^f_w = \sum_{x\notin\mathcal{D}} c^f_x.$$

Assumption 2 (All new technologies adopted) *For any vector of wages with $w_x < \bar{w}$ for $x \in \mathcal{D}$ and a wage $w > \bar{w}$ in undisrupted islands such that $c^f(\{w_x\}_{x\in\mathcal{D}}, w) = 1$, we have*

$$\frac{c^f_x(\{w_x\}_{x\in\mathcal{D}}, w)}{c^f_w(\{w_x\}_{x\in\mathcal{D}}, w)} > \frac{c^f_x(\{\bar{w}\}_{x\in\mathcal{D}}, \bar{w})}{c^f_w(\{\bar{w}\}_{x\in\mathcal{D}}, \bar{w})}$$

This assumption ensures that new technologies are adopted in all disrupted islands (so long as the after-tax cost of the new technology is below the initial market wage). The assumption prevents adoption in one island from reducing the relative demand for goods produced in other disrupted island.

Assumptions 1 and 2 hold when there is a single disrupted and a single undisrupted island, but also when there are many islands whose outputs are combined via a constant-elasticity of substitution aggregator, f.

The following propositions characterize the transitional dynamics of the economy in terms of wages and employment across islands.

Proposition 1 *Suppose that Assumptions 1 and 2 hold and that $\bar{w} > (1+\tau_{x,t})/A_{x,t}$, so that the new technologies are adopted at time 0. Wages are given by*

$$w_{x,t} = \begin{cases} (1+\tau_{x,t})/A_{x,t} & \text{if } x \in \mathcal{D} \\ w_t & \text{if } x \notin \mathcal{D}, \end{cases}$$

where the common wage w_t in undisrupted islands satisfies

$$1 = c^f(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t).$$

Along the transition, island employment is given by $\ell_{x,t} = e^{-\alpha_x t} \cdot \ell_{x,0}$ for $x \in \mathcal{D}$ and $\ell_t =$

$1 - \sum_{x \in \mathcal{D}} e^{-\alpha_x t} \cdot \ell_{x,0}$ *for the remaining islands; output is given by*

$$y_t = \ell_t \cdot \frac{1}{c_w^f\left(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t\right)};$$

and new technology utilization at island $x \in \mathcal{D}$ is given by

$$k_{x,t} = \ell_t \cdot \frac{c_x^f\left(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t\right)}{c_w^f\left(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t\right)} - \ell_{x,t} > 0.$$

Proposition 2 *Suppose Assumptions 1 and 2 hold.*

- *If households are insured against the island disruption, and there is risk sharing inside islands, $U_{x,0} = U_0$ and $c_{x,t} = c_t$ across all islands.*
- *If households had the same assets before the shock and they are not insured against the disruption, then $a_{x,0} = a_0$ and $U_{x,0} < U_0$.*
- *If households are hand-to-mouth, then $U_{x,0} < U_0$.*

2 Optimal Policy and the Gains from Gradualism

We maintain Assumptions 1 and 2 and consider the case where households are not insured against the disruption, so that there is a role for policy.

We evaluate policies using a symmetric welfare function $W_0 = \sum_{x\in\mathcal{X}} \int_{h\in x} \mathcal{W}\left(U_{x,0}^h\right) \cdot dh$, where $U_{x,0}^h = U_{x,0}$ is the expected lifetime utility of household h in island x after the shock and $\mathcal{W}$ is an increasing and concave function. The per-capita Pareto weights (or social marginal welfare weights) for households from island x are therefore

$$g := \mathcal{W}'(U_0) \geq 0 \text{ for } x \notin \mathcal{D} \qquad\qquad g_x := \mathcal{W}'(U_{x,0}) \geq 0 \text{ for } x \in \mathcal{D}.$$

This implies that $g_x \geq g$ capturing the incentives to compensate losers from the disruption.[10]

Let $\mathcal{U}$ and $\mathcal{U}_x$ denote indirect utilities for households who were initially at undisrupted

[10]Our formulas for optimal taxes apply more generally for Pareto weights capturing asymmetries across islands, or when using *generalized social marginal welfare weights* that depend on broader ethical and political considerations (as in Saez and Stantcheva, 2016). Both features can be captured by having g_x and g depend on additional arguments. To emphasize the generality of our formulas, we leave the arguments of g_x and g unspecified and treat them as general functions of the allocation of resources across islands.

islands and disrupted island x, respectively:

$$U_0 = \mathcal{U}(\{w_t + T_t\}_{t=0}^{\infty}), \qquad U_{x,0} = \mathcal{U}_x(\{w_{x,t} + T_t, w_t + T_t\}_{t=0}^{\infty}; \alpha_x) - \kappa_x(\alpha_x).$$

Of all households from island $x \in \mathcal{D}$, a fraction $P_{x,t} = e^{-\alpha_x \cdot t}$ will still work in the disrupted island at time t and consume $c_{x,t}$. The remaining $1 - P_{x,t}$ households will have reallocated by that time, with a fraction $\alpha_x \cdot e^{-\alpha_x \cdot t_r}$ reallocating at time $t_r \in [0, t]$. We denote their consumption at time t by $c_{x,t_r,t}$. Households at undisrupted islands do not reallocate and face no uncertainty. We denote their consumption at time t by c_t.

We first provide a general lemma that characterizes the change in welfare resulting from a variation in taxes and technology utilization. This lemma relates to Lemma 1 in Costinot and Werning (forthcoming), but differs in that it accounts for the fact that variations in taxes affect households' incomes at all future states. It also builds on variational arguments from Saez (2001); Saez and Stantcheva (2016); Tsyvinski and Werquin (2017).

Lemma 1 (Variations Lemma) *Consider a variation in taxes that induces a change in wages $dw_t, dw_{x,t}$, technology $dk_{x,t}$, tax revenue dT_t, and reallocation effort $d\alpha_x$. This variation changes tax revenue by*

$$dT_t = -\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot dw_{x,t} - \ell_t \cdot dw_t + \sum_x \tau_x \cdot \frac{dk_{x,t}}{A_{x,t}} \tag{1}$$

and social welfare by

$$dW_0 = \int_0^{\infty} \left[\sum_{x \in \mathcal{D}} \ell_{x,0} \cdot g_x \cdot \Big(P_{x,t} \cdot \lambda_{x,d,t} \cdot (dw_{x,t} + dT_t) + (1 - P_{x,t}) \cdot \lambda_{x,r,t} \cdot (dw_t + dT_t) \Big) + \ell_0 \cdot g \cdot \lambda_t \cdot (dw_t + dT_t) \right] \cdot dt, \tag{2}$$

where $\lambda_{x,d,t} = e^{-\rho t} \cdot u'(c_{x,t})$ denotes the marginal utility of consumption at time t for households that have not reallocated, $\lambda_{x,r,t} = \mathbb{E}[e^{-\rho t} \cdot u'(c_{x,t_r,t}) | t_r \leq t]$ denotes the average marginal utility of consumption among households that reallocated by time t, and $\lambda_t = e^{-\rho t} \cdot u'(c_t)$ denotes the marginal utility of consumption of non-disrupted households at time t.

The lemma shows that the benefits of distorting technology depend on whether this redistributes income towards households with a high marginal utility of consumption.[11] In

[11]An important assumption in our model is that all households are ex-ante equal. This deliberate

practice, this is not observed; most papers provide estimates of the income drop experienced by households in disrupted islands. When taking our model to the data, our approach is to calibrate the model to match the income decline estimated in Autor et al. (2014) and Cortes (2016), and infer marginal utilities of consumption by considering four scenarios that summarize households' consumption and saving behavior:

1. Hand-to-mouth (transition risk and no borrowing): this scenario assumes households are hand-to-mouth. This implies $\lambda_{x,d,t} = e^{-\rho t} \cdot u'(w_{x,t} + T_t)$, $\lambda_{x,r,t} = e^{-\rho t} \cdot u'(w_t + T_t)$, and $\lambda_t = e^{-\rho t} \cdot u'(w_t + T_t)$. In this scenario, households cannot borrow to smooth their consumption along the transition and face the risk of transitioning late.

2. No borrowing and no transition risk: this scenario assumes no borrowing from other islands or foreigners. However, we allow households to engage in risk sharing within their island to share the risks of transitioning late. Equivalently, one could think of this as a case where each household owns a mass 1 of units of labor, which it then retools at a rate α to be used in other islands, so that it faces no uncertainty. This implies $\lambda_{x,d,t} = \lambda_{x,r,t} = e^{-\rho t} \cdot u'(P_{x,t} \cdot w_{x,t} + (1 - P_{x,t}) \cdot w_t + T_t)$, and $\lambda_t = e^{-\rho t} \cdot u'(w_t + T_t)$. In this scenario, households cannot borrow to smooth their consumption along the transition but do not face the risk of transitioning late.

3. Borrowing with transition risk: this scenario assumes that households can borrow at exogenous interest rate r but face the risk of transitioning late. In disrupted islands, households problem can be summarized by the following system of HJB equations

$$\rho v_x(a, t; \alpha) - \dot{v}_x(a, t; \alpha) = \max_c u(c) + \partial_a v_x(a, t; \alpha) \cdot (ra + w_{x,t} - c) + \alpha_x \cdot (v(a, t) - v_x(a, t; \alpha)),$$
$$\rho v(a, t) - \dot{v}(a, t) = \max_c u(c) + \partial_a v(a, t) \cdot (ra + w_t - c).$$

Here, $v_x(a, t; \alpha)$ is the value function of households in disrupted islands at time t with assets a when they exert reallocation effort α, and $v(a, t)$ is the value function of households in undisrupted islands with assets a. This problem can be solved numerically using the tools from Achdou et al. (2021), and produces paths for $\lambda_{x,d,t} = e^{-\rho t} \cdot u'(c_{x,t})$ for households in disrupted islands, paths for $\lambda_{x,r,t} = \mathbb{E}[e^{-\rho t} \cdot u'(c_{x,t_r,t}) | t_m \leq t]$ for households that reallocate, and paths for $\lambda_t = e^{-\rho t} \cdot u'(c_t)$ as functions of the stream of incomes in disrupted and

choice allowed us to focus on the problem of how to tax new technologies so as to protect losers along the transition. The Appendix extends our optimal tax formulas to an environment with inequality across households within and between islands. We show that, under mild assumptions, inequality within islands does not affect our optimal tax formulas. On the other hand, inequality across islands creates incentives for stronger responses if disrupted islands had lower incomes (or fewer assets) to begin with.

undisrupted islands, the reallocation rate α, and initial asset holding. The details on how to solve this problem are provided in Appendix A.5.

4. Borrowing with no transition risk: this scenario assumes ex-post complete markets. That is, households can freely save and borrow at exogenous interest rate r and share transition risks within their islands. Households' problem becomes

$$\max \int_0^\infty e^{-\rho t} \cdot u(c_{x,t}) \cdot dt \quad \text{s.t.: } 0 \leq \int_0^\infty e^{-rt} \cdot \left[P_{x,t} \cdot w_{x,t} + (1 - P_{x,t}) \cdot w_t + T_t - c_{x,t}\right] \cdot dt + a_{x,0}.$$

which implies $\lambda_{x,d,t} = \lambda_{x,r,t} = e^{-rt} \cdot u'(c_{x,0})$ and $\lambda_t = e^{-rt} \cdot u'(c_0)$.

2.1 Optimal Policy with Exogenous Reallocation

The following Proposition provides our first formula for optimal taxes.

Proposition 3 *Let χ_t be the per-capita marginal social value of increasing income at undisrupted jobs at time t and $\chi_{x,t} \geq \chi_t$ be the per-capita social value of increasing income in island x at time t, and $\bar{\chi}_t$ their population-weighted average. With exogenous reallocation effort, a necessary condition for an optimal tax sequence is that*

$$\tau_{x',t} = \sum_{x \in \mathcal{D}} \frac{\ell_{x,t} \cdot w_{x,t}}{m_{x',t}} \cdot \left(\frac{\chi_{x,t}}{\bar{\chi}_t} - 1\right) \cdot \left(-\frac{\partial \ln w_{x,t}}{\partial \ln k_{x',t}}\right) + \frac{\ell_t \cdot w_t}{m_{x',t}} \cdot \left(\frac{\chi_t}{\bar{\chi}_t} - 1\right) \cdot \left(-\frac{\partial \ln w_t}{\partial \ln k_{x',t}}\right), \tag{3}$$

where $m_{x,t}$ denotes expenditure on $k_{x,t}$, and the χ's are given by

$$\chi_{x,t} = g_x \cdot \lambda_{x,d,t} \qquad\qquad \chi_t = \frac{1}{\ell_t} \cdot \left(\sum_{x \in \mathcal{D}} \ell_{x,0} \cdot (1 - P_{x,t}) \cdot g_x \cdot \lambda_{x,r,t} + \ell_0 \cdot g \cdot \lambda_t\right).$$

Moreover, the formula in equation (3) implies that $\lim_{t\to\infty} \tau_{x',t} = 0$.

The derivation follows Costinot and Werning (forthcoming). The key idea is that, at an optimum, a variation that changes the quantity of new technology used in island $x' \in \mathcal{D}$ at time t should not change welfare. Using Lemma 1 to evaluate this variation yields the optimal tax formula in equation (3). One can think of the left-hand side of (3) as the marginal benefit of increasing $k_{x',t}$—the additional tax revenue—and the right-hand side as the marginal cost—the adverse distributional implications along the transition.

The optimal tax on $k_{x',t}$ depend on three factors:

1. Taxes should be higher when an exogenous increase in $k_{x',t}$ has a sizable negative im-

pact on the wage of disrupted islands at time t, as captured by the partial derivatives $-\partial \ln w_{x,t}/\partial k_{x',t}$. In this case, reducing the use of the new technology via taxes is an effective tool to redistribute income towards households in disrupted islands.

2. Taxes should be higher when households in disrupted islands have a high marginal utility of consumption during the transition, as captured by the marginal social values $\chi_{x,t} \geq \chi_t$. This depends on the income decline experienced by disrupted households, which is informed by the evidence in Cortes (2016) for routine jobs and Autor et al. (2014) for the China Shock, and households' ability to mitigate this drop in income by relying on past savings and borrowing.

3. Taxes should be lower in the long run. This is because a variation that increases income at disrupted islands in the distant future has a small effect on the welfare of households in that island today, since they expect to reallocate and (in some of our scenarios) cannot borrow against that future income. In fact, the optimal long-run tax is zero since $\ell_{x,t} \to 0$ and all workers reallocate away from disrupted islands.

2.2 Endogenous Reallocation

We now extent Proposition 3 to the case with endogenous reallocation. With endogenous reallocation, a variation in policy that affects wages $w_{x,t}$ and w_t in the future necessarily changes reallocation rates $d\alpha_x$ at island x, since reallocation is a forward looking decision. The effect of $d\alpha_x$ on transition probabilities is second order, because households internalize this benefit. However, $d\alpha_x$ brings general equilibrium effects on factor prices (for a given level of capital utilization) and revenue that might be beneficial for social welfare.

To account for these effects, we need additional notation. Let $\mathcal{U}_{x,\alpha} = \partial_\alpha \mathcal{U}_x$ denote the utility gains of changing the reallocation rate at the margin for a household in island x. Households' choice of α satisfies the first-order condition $\mathcal{U}_{x,\alpha} = \kappa'_x(\alpha)$. Let $\mathcal{U}_{x,\alpha,d,t} \cdot dt$ denote the marginal effect of changes in income at time t in island x on $\mathcal{U}_{x,\alpha}$ and $\mathcal{U}_{x,\alpha,r,t} \cdot dt$ denote the marginal effect of changes in income at time t in non-disrupted islands on $\mathcal{U}_{x,\alpha}$. The marginal change $\mathcal{U}_{x,\alpha,d,t}$ is typically negative and $\mathcal{U}_{x,\alpha,r,t}$ positive, reflecting the disincentives of policies that tax new technology on reallocation efforts. Define $\varepsilon_{x'',x} \cdot \epsilon$ as the rate of change in $\alpha_{x''}$ when $\mathcal{U}_{x,\alpha}$ changes by ϵ. The cross partials $\varepsilon_{x'',x}$ depend on the curvature of the cost function $\kappa_x(\alpha_x)$ and the way in which reallocation away from island x affects factor prices and tax revenue, shaping incentives for reallocating away from island x''.

When $\varepsilon_{x'',x} = 0$ for all x'', x we are back in the exogenous reallocation case.[12]

Proposition 4 provides formulas for optimal taxes in terms of $\mathcal{U}_{x,\alpha}$, $\mathcal{U}_{x,\alpha,d,t}$, $\mathcal{U}_{x,\alpha,r,t}$, and $\varepsilon_{x,x''}$. The Appendix provides explicit formulas for these objects in terms of primitives.

PROPOSITION 4 *When effort is endogenous, a necessary condition for an optimal tax sequence is that*

$$\tau_{x',t} = \sum_{x\in\mathcal{D}} \frac{\ell_{x,t}\cdot w_{x,t}}{m_{x',t}} \cdot \left(\frac{\chi^{end}_{x,t}}{\bar{\chi}^{end}_t} - 1\right) \cdot \left(-\frac{\partial \ln w_{x,t}}{\partial \ln k_{x',t}}\right) + \frac{\ell_t \cdot w_t}{m_{x',t}} \cdot \left(\frac{\chi^{end}_t}{\bar{\chi}^{end}_t} - 1\right) \cdot \left(-\frac{\partial \ln w_t}{\partial \ln k_{x',t}}\right), \tag{4}$$

where the χ^{end}'s are now given by

$$\chi^{end}_{x,t} = \chi_{x,t} + \sum_{x''\in\mathcal{D}} \frac{\ell_{x'',0}}{\ell_{x,t}} \cdot \mu_{x''} \cdot \varepsilon_{x'',x} \cdot \mathcal{U}_{x,\alpha,d,t}, \qquad \chi^{end}_t = \chi_t + \sum_{x,x''\in\mathcal{D}} \frac{\ell_{x'',0}}{\ell_t} \cdot \mu_{x''} \cdot \varepsilon_{x'',x} \cdot \mathcal{U}_{x,\alpha,r,t}.$$

Here, μ_x is the social value of increasing the reallocation rate per displaced worker holding all other quantities constant:

$$\mu_x = \int_0^\infty (-s \cdot e^{-\alpha_x s}) \cdot \left[\sum_{x''\in\mathcal{D}} \ell_{x'',s} \cdot (\chi_{x'',s} - \bar{\chi}_s) \cdot \frac{\partial w_{x'',s}}{\partial \ell_{x,s}} + \ell_s \cdot (\chi_s - \bar{\chi}_s) \cdot \frac{\partial w_s}{\partial \ell_{x,s}}\right] ds. \tag{5}$$

The formula for optimal taxes shares the same structure as before. All that is needed is redefining the χ's, so that now the social marginal value of increasing future income at different islands accounts for their effect on reallocation rates and the social benefit of reallocation (namely, higher wages at disrupted islands and higher tax revenues).[13]

One important implication of the formula is that, in some relevant cases, it justifies a short-run tax on new technologies and trade accompanied by a subsidy in the medium run.

[12] This specification assumes that the cost of reallocating is independent of the number of people doing so. The Appendix provides results for a case where endogenous reallocation efforts lead to congestion. For example, retraining programs might face some form of decreasing returns to scale, making it easier to retrain a small fraction of workers each period than a large fraction immediately. Congestion provides another rationale for gradualism, even in the absence of distributional concerns.

[13] An alternative policy tool entails subsidizing reallocation efforts, for example via retraining subsidies or active labor-market policies. To understand how the availability of these policies affects optimal taxes, the Appendix considers a case in which the planner controls α_x directly and sets it at its socially optimal level. When subsidies to reallocation are available, we are back to the same formula for optimal taxes derived for the case with exogenous effort in Proposition 3. This does not mean that the optimal tax path is the same as before: the multipliers on the right-hand side are now evaluated along an equilibrium with the socially optimal level of α_x, which involves more rapid reallocation. The correct interpretation is that, in this case, concerns about a more gradual transition dampening reallocation efforts should not factor in the decision of how to tax new technologies.

For example, when households are hand to mouth, we have

$$
\begin{aligned}
\mathcal{U}_{x,\alpha} &= \int_0^\infty e^{-\rho t} \cdot (t \cdot P_{x,t}) \cdot [u(w_t + T_t) - u(w_{x,t} + T_t)] \cdot dt, \\
\mathcal{U}_{x,\alpha,d,t} &= -(t \cdot P_{x,t}) \cdot \lambda_{x,d,t}, \\
\mathcal{U}_{x,\alpha,r,t} &= (t \cdot P_{x,t}) \cdot \lambda_{x,r,t}.
\end{aligned}
$$

Social marginal values of increasing incomes at disrupted and other islands become

$$
\begin{aligned}
\chi^{\text{end}}_{x,t} &= \chi_{x,t} - \sum_{x'' \in \mathcal{D}} \frac{\ell_{x'',0}}{\ell_{x,t}} \cdot \mu_{x''} \cdot \varepsilon_{x'',x} \cdot (t \cdot P_{x,t}) \cdot \lambda_{x,d,t}, \\
\chi^{\text{end}}_{t} &= \chi_{t} + \sum_{x, x'' \in \mathcal{D}} \frac{\ell_{x'',0}}{\ell_{t}} \cdot \mu_{x''} \cdot \varepsilon_{x'',x} \cdot (t \cdot P_{x,t}) \cdot \lambda_{x,r,t}.
\end{aligned}
$$

For small t, we always have $\chi^{\text{end}}_{x,t} > \chi^{\text{end}}_{t}$, and short-run taxes are optimal. For intermediate t we could have $\chi^{\text{end}}_{x,t} < 0$ and a subsidy to $k_{x,t}$ would be justified at that point. The short run taxes are useful for protecting losers. The medium run subsidy provides back-loaded incentives for reallocation. As before, a zero tax is optimal in the long run, since $t \cdot P_{x,t} \to 0$.

2.3 Assistance Programs and Other Policy Tools

The optimal tax on new technologies depends on the availability of alternative tax instruments. If island-specific transfers, wage subsidies, or assistance programs were available, there would be no rationale for distorting technology—a consequence of the Second Welfare Theorem and Diamond and Mirrlees (1971). However, these alternative (and more desirable) instruments are limited in practice, since identifying and targeting losers from trade and technological change is challenging.[14]

Following Naito (1999) and Costinot and Werning (forthcoming), we focus here on a more realistic set of tax instruments and assistance programs that depend only on households' income history. Appendix A.2.2 provides the details of our formulation. We consider the role of temporary assistance programs that work by replacing a fraction $\mathcal{R}_t$ of the drop in income experienced by workers that enter the program relative to some past or reference level. To capture the limitations of these programs, we allow households to affect their income by making endogenous decisions over work effort, $n_{x,t}$. For example, households

[14] There is also the possibility that reforming the tax system to deal with short-run disruption costs might run into other problems. For example, it might be politically costly to change the structure of the income tax system temporarily in an effort to compensate losers, since there is a risk that these reforms will become entrenched.

may reduce their work effort and get fired because of shirking and not because of technological disruptions. Or households may decide to stop investing in their skills, causing their income to drop, not because of technology but due to their choices. We then model assistance programs as transfers that depend solely on current income, status quo income, and the replacement rate $\mathcal{R}_t$:

$$\text{assistance program transfer}_t^h = \mathcal{R}_t \cdot \left(y_0^h \cdot \frac{w_t \cdot n_t^*}{\bar{w}} - y_t^h \right).$$

Here, $y_0^h \cdot \frac{w_t \cdot n_t^*}{\bar{w}}$ is the expected income path of a household unaffected by trade, technology, or trade liberalization and y_t^h is realized income. In this expression, n_t^* is the level of work effort by an undisrupted household that does not participate in the program. The assistance program then compensates participants for a fraction $\mathcal{R}_t$ of their income loss, where $\mathcal{R}_t \in [0,1]$ is the replacement rate of the program.[15]

The next proposition characterizes optimal policy when governments can tax new technologies and set temporary assistance programs characterized by a path $\mathcal{R}_t$.

Proposition 5 *Suppose that work effort is not affected by income effects and responds to wages with an elasticity $\varepsilon_\ell \geq 0$. Suppose also that the reallocation rate is exogenous. The optimal taxes on new technologies $\{\tau_{x,t}\}$ and path for assistance programs $\{\mathcal{R}_t\}$ satisfy the necessary conditions*

$$(6) \qquad \tau_{x',t} = \sum_{x \in \mathcal{D}} \frac{\ell_{x,t} \cdot n_{x,t} \cdot w_{x,t}}{m_{x',t}} \cdot (1 - \mathcal{R}_t) \cdot \left(\frac{\chi_{x,t}}{\bar{\chi}_t} - 1 \right) \cdot \left(- \frac{\partial \ln w_{x,t}}{\partial \ln k_{x',t}} \right) + \frac{\ell_t \cdot n_t \cdot w_t}{m_{x',t}} \cdot (1 - \mathcal{R}_t) \cdot \left(\frac{\chi_t}{\bar{\chi}_t} - 1 \right) \cdot \left(- \frac{\partial \ln w_t}{\partial \ln k_{x',t}} \right) + \mathcal{R}_t \cdot \varepsilon_\ell \cdot \frac{d \ln w}{d \ln k_{x',t}},$$

and

$$(7) \qquad \frac{\mathcal{R}_t}{1 - \mathcal{R}_t} = \frac{1}{\varepsilon_\ell} \cdot \frac{\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot n_{x,t} \cdot w_{x,t} \cdot \left(1 - \frac{\chi_{x,t}}{\bar{\chi}_t} \right) + \ell_t \cdot n_t \cdot w_t \cdot \left(1 - \frac{\chi_t}{\bar{\chi}_t} \right)}{\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot n_{x,t} \cdot w_{x,t} + \ell_t \cdot n_t \cdot w_t} + \frac{1}{\varepsilon_\ell} \cdot (1 - \mathcal{R}_t) \cdot \frac{\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot n_{x,t} \cdot \left(\frac{\chi_{x,t}}{\bar{\chi}_t} - 1 \right) \cdot \frac{d w_{x,t}}{d \mathcal{R}_t} + \ell_t \cdot n_t \cdot \left(\frac{\chi_t}{\bar{\chi}_t} - 1 \right) \cdot \frac{d w_t}{d \mathcal{R}_t}}{\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot n_{x,t} \cdot w_{x,t} + \ell_t \cdot n_t \cdot w_t}.$$

[15] Appendix A.2.2 shows that these programs are equivalent to a time-varying linear tax on income at a rate $\mathcal{R}_t$, so one can also think of the path for $\mathcal{R}_t$ as describing an optimal temporary income tax reform. This equivalence breaks down when there are workers with different initial income levels, as in the extensions in Appendix A.4.1. In practice, one can think of assistance programs as including unemployment and disability insurance, which work by replacing part of workers' income loss as they become unemployed and claim disability benefits. Although these programs condition on workers' being out of a job, but cannot distinguish if this happened for exogenous or endogenous reasons.

Here, $\frac{dw_t}{d\mathcal{R}_t}$ and $\frac{dw_{x,t}}{d\mathcal{R}_t}$ give the aggregate change in wages resulting from a change in the replacement rate and its effects on work effort, and $\frac{d\ln w}{d\ln k_{x',t}}$ gives the effect of changing capital utilization on average wages.

The proposition shows that, when $\varepsilon_\ell = 0$, it is optimal to set assistance programs with a full replacement rate. In this case, we do not need to tax new technologies. However, in the more relevant case where $\varepsilon_\ell > 0$, it is optimal to tax technology in the short run. The proposition also shows that, as ε_ℓ increases, so that assistance programs are exploited more frequently, the optimal replacement rate will be lower and higher short-run taxes on technology are justified. Taxing new technologies is desirable because it assists workers affected by exogenous disruptions and not those who reduced their work effort to exploit assistance programs.[16]

2.4 Optimal Timing of Reforms

We can also apply the formulas in Propositions 3, 4 and A3 to the question of how to time economic reforms. Consider a variant of our model where the new technology already existed at time 0, had constant productivity A_x over time, and was not in use because there was a distortionary tax $\bar{\tau}_{x,0} > 0$ in place preventing this, so that

$$\frac{1+\bar{\tau}_{x,0}}{A_{x,0}} = \bar{w} \text{ for } x \in \mathcal{D}.$$

Suppose that at time 0, policymakers decide to pursue a reform that removes this distortion. The propositions characterize the optimal path for taxes, which involves an instantaneous jump to $\tau_{x,0} \in (0, \bar{\tau}_{x,0})$ and from there on a gradual decline to $\tau_{x,t} = 0$.

2.5 The Gains from Technological Gradualism

The previous section showed that it is optimal to tax technology in the short run to assist disrupted households. That does not necessarily mean that the economy would benefit from moving to a counterfactual world where technological progress takes place more gradually. These are different questions because they consider different perturbations of the economy. Reducing the use of a new technology via taxes generates a more gradual path for wages

[16] This is the same rationale for why it is optimal to distort technology in Naito (1999) and Costinot and Werning (forthcoming). For example, in Costinot and Werning (forthcoming), distorting technology is desirable because it redistributes towards workers whose wages are low because of technological reasons, and not towards those whose income is low due to lack of work effort.

for disrupted workers *and* raises tax revenue to be distributed among households. Instead, moving to a counterfactual world where technological progress happens more gradually shifts the production possibility frontier and does not generate additional tax revenue.

To answer this different question, we compute *the gains from technological gradualism.*[17] Suppose that the observed path for technology is differentiable with respect to time and given by $\{A_{x,t}\}_{x \in \mathcal{D}}$. The gains from technological gradualism capture the welfare gains of facing a counterfactual path of technology given by $\{A_{x,(1-\Gamma)\cdot t}\}_{x \in \mathcal{D}}$ with Γ determining the level of gradualism. In this alternative path, technology progresses more gradually $(\Gamma > 0)$ or more rapidly $(\Gamma > 0)$ but converges to the same steady state level.

Proposition 6 *Suppose that the government taxes new technologies optimally. The welfare gains from technological gradualism are always negative and more rapid technological change raises welfare. On the other hand, if the government does not tax new technologies nor sets assistance programs, the welfare gains from gradualism around $\Gamma \approx 0$ have an ambiguous sign and are*

$$\mathcal{W}_g = \frac{\partial W_0}{\partial g} = \int_0^\infty \left\{ \sum_{x \in \mathcal{D}} \left(\ell_{x,t} \cdot \chi_{x,t} \cdot w_{x,t} - \ell_t \cdot \chi_t \cdot w_t \cdot \frac{s_{x,t}}{s_t} \right) \cdot \frac{\dot{A}_{x,t}}{A_{x,t}} \cdot t \right\} \cdot dt, \tag{8}$$

with $s_{x,t}$ the share of island x and $s_t = 1 - \sum_{x \in \mathcal{D}} s_{x,t}$ the share of undisrupted islands in GDP.

The Proposition shows that the gains from technological gradualism depend on tax policy. When optimal taxes on new technologies are in place, more rapid technological progress is always welcomed. This holds even if the government has no other tools for redistribution or assistance programs. This result follows from a simple envelop logic: a government can always mimic a less gradual path for technology by taxing it. In the worst case scenario, the government can respond to more rapid technological change by raising taxes so as to keep wages and allocations unchanged while at the same time increasing tax revenue.[18]

[17] An active literature in trade quantifies the gains from trade—how much welfare would decrease if a country remained in autarky instead of engaging in the observed level of trade. See for example Costinot and Rodríguez-Clare (2014) for a review of the literature quantifying the gains from trade, and Antràs et al. (2017) and Galle et al. (2022) for an approach to extend these welfare gains formulas to an environment with inequality. See also Caliendo et al. (2019) for work quantifying the gains from trade accounting for sluggish transitional dynamics. We can think of the gains from trade or technology in our context as the change in welfare resulting from the arrival of the new technology. The gains from gradualism, on the other hand, capture how the gains from trade vary with the graduality of the shock.

[18] This is the same argument used in the working paper version of Costinot and Werning (2018) to show that, with optimal taxes in place, technology always raises welfare and should be valued in the same way as in a first best world.

When the government takes no action, the gains from technological gradualism can be positive or negative. More rapid technological progress will be a positive force independently of government policy, if: i. differences in consumption between disrupted and non-disrupted households are small; ii. workers reallocate rapidly; and iii. technological advances are back-loaded, so that technology increases slowly initially. When these conditions are not met, and governments do not intervene, the gains from technological gradualism are positive and society benefits from slower technological progress.[19]

3 Application I: The Automation of Routine Jobs

3.1 Description, Empirical Evidence, and Calibration

Description: There are 5 islands. Islands 2–5 are in $\mathcal{D}$ and represent segments of routine occupations $o(x)$ (where $o(x)$ denotes the occupation associated with island x) disrupted by technological progress: i. clerical and administrative occupations (10% of employment in 1980); ii. sales occupations (5% of employment in 1985); iii. production occupations (18.5% of employment in 1980); iv. transportation and material handling occupations (4% of employment in 1980). These four occupational groups are identified as routine in Acemoglu and Autor (2011). Their wage and reallocation dynamics are documented in Cortes (2016). The first island represents segments of these occupations that have not been affected by the automation of routine jobs plus non-routine occupations.

Not all jobs that are part of an occupation are replaced by technology. Only a fraction $s_{o(x)}$ of all jobs in occupation $o(x)$ are disrupted and belong to island x. This implies that island x accounts for a fraction $s_{o(x)} \cdot \Omega_{o(x)}$ of initial employment, where $\Omega_{o(x)}$ denotes the initial share of employment in occupation $o(x)$.[20] We treat $s_{o(x)} \in [0, 1]$ as an unobservable to be calibrated in order to match the scope of the technological disruption.

The empirical evidence in Cortes (2016): This paper estimates trends in occupational wages over time and provides evidence that workers initially employed in routine jobs

[19]The formula for the gains from gradualism holds both for endogenous and exogenous effort. The reason is that changes in reallocation effort in response to more gradual technological advances "envelope out." This is why the formulas only depend on marginal social values of consumption, $\chi_{x,t}$ and χ_t, and not on the change in reallocation effort.

[20]One may consider a mass 1 of islands partitioned into occupations. Each island represents a differentiated job within an occupation, with island x belonging to occupation $o(x)$, and a mass $\Omega_{o(x)}$ of islands in each occupation. The automation of routine jobs corresponds to an improvement in the productivity of specialized capital, equipment, and software that leads to the automation of a share $s_{o(x)}$ of the islands that are part of occupation $o(x)$.

experienced lower future wage growth. Cortes uses data from the Panel Study of Income Dynamics (PSID) to estimate a variant of the model:

$$\log \text{hourly wage}_{i,o,t} = \delta_t + \text{Routine}_o \cdot \theta_t + X'_{i,t} \cdot \zeta + X'_i \cdot \zeta_t + \gamma_{i,o} + u_{i,o,t}. \tag{9}$$

The model explains trends in hourly wages for individual i employed in occupation o at time t as a function of: i. common time trends, δ_t; ii. a differential time-path for routine occupations, $\text{Routine}_o \cdot \theta_t$, where Routine_o is a dummy that takes the value of 1 for routine occupations and θ_t captures differential wage trends in these jobs over time; iii. time varying individual covariates $X'_{i,t} \cdot \zeta$; iv. differential time trends by individual fixed characteristics $X'_i \cdot \zeta_t$; and v. permanent differences in the wage of individual i in occupation o, $\gamma_{i,o}$. The last term accounts for selection in persistent attributes that make some individuals more productive at some occupations.[21]

Panels A and B in Figure 1 provide estimates of equation (9). Panel A reports estimates of θ_t from Cortes' data for different specifications: 1. controlling for permanent wage differences by individuals across occupations and demographics; 2. allowing for differential trends by region of residence and whether the person resides in urban or suburban areas; 3. controlling for union membership; and 4. allowing for differential trends over time by education level. All specifications show a permanent decline of 25–30% in relative wages paid for routine jobs since 1986. The more demanding specification that controls for trends in wages by educational levels dates the start of the decline to 1986. Panel B reports separate estimates $\theta_{o(x),t}$ for different routine occupations. All routine occupations exhibit a similar pattern of declining wages, though the speed and extent of the decline varies.

Cortes (2016) estimates of occupation wages over time are informative of the behavior of $w_{x,t}$ in our model, which in turn provides information on the advances in technologies capable of substituting workers in some routine jobs, $A_{x,t}$. For example, if all jobs within an occupation are disrupted, then $\ln(w_{o(x),t}/w_t) = \theta_{o(x),t}$, which one can use to invert for advances in technology $A_{x,t}$.

Cortes (2016) also provides estimates of future wage growth for workers employed in routine occupations at time t_0. In particular, Cortes estimates:

$$\Delta \log \text{wage income}_{i,o,t} = \delta_t + \beta \cdot \text{Routine}_{o,t_0} + X'_{i,t_0} \cdot \zeta + u_{i,o,t}. \tag{10}$$

[21] Cortes (2016) also estimates the price associated with cognitive occupations, but for our purposes, the relevant object is the price of routine occupations relative to all others.

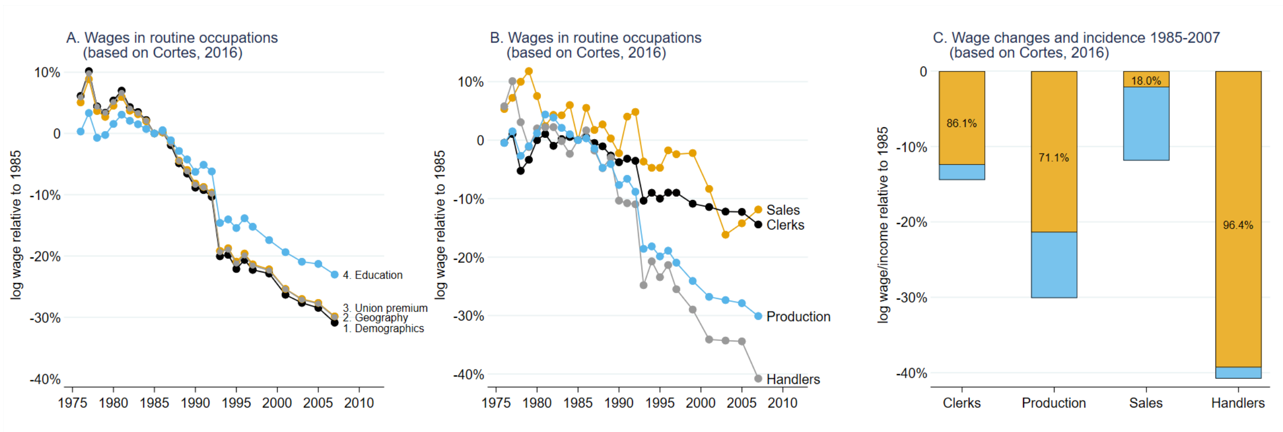

FIGURE 1: ESTIMATES OF WAGE TRENDS FOR ROUTINE OCCUPATIONS. Panel A reports estimates of θ_t in equation (9) using the data and specification from Cortes (2016). Panel B reports separate estimates of $\theta_{o(x),t}$ by occupational group. Panel C reports estimates of the incidence of these shocks on workers that held routine occupations in 1985.

This model explains wage growth between period t_0 and t as a function of individual characteristics at time t_0 and a dummy for whether the individual worked at a routine occupation at time t_0.[22] We use Cortes data an estimate this regression for $t_0 = 1985$ and $t = 2007$, controlling for age, demographics, union membership and education in 1985. We estimate that individuals in routine jobs experienced 16%–20% less wage growth over this period than comparable workers, depending on the set of covariates used, which aligns with the 20-year growth estimates from Table 2 in Cortes (2016). These estimates are informative of the incidence of shocks reducing routine wages over time on workers that held these jobs initially.

Panel C in Figure 1 reports separate estimates by occupation for the occupational wage changes $\theta_{o(x),t}$ from 1985 to 2007 (blue bar) and the incidence of this shock on workers who held these occupations in 1985. Except for sales jobs, there is a large incidence of the shock, with workers who held clerical, production, or handling jobs by 1985 experiencing a wage decline of 71%–96% the amount that they would have experienced if they were not able to reallocate and had hold to their initial jobs. On average, the incidence of the shock is of 71%, which shows that reallocation helps workers mitigate some (but not all) of the adverse impacts of the shock over time. Through the lens of our model, the high incidence estimated by Cortes (2016) points to small reallocation rates α_x out of disrupted jobs.

[22]We focus on a variant of the specification used in Cortes (2016) that looks at totla income accounting for hours worked. This is because reduced work hours might be an important margin of adjustment for workers. This strategy does not account for non-employment, though we did not find evidence of sizable employment effects in the PSID.

Calibration: Aggregate output y_t is given by a CES aggregator

$$y_t = \left(\nu^{\frac{1}{\sigma}} \cdot \ell_t^{\frac{\sigma-1}{\sigma}} + \sum_{x \in \mathcal{D}} \nu_x^{\frac{1}{\sigma}} \cdot y_{x,t}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{11}$$

We take a value of $\sigma = 0.85$ from the literature on occupational polarization (see Goos et al., 2014). We normalize status-quo wages to $\bar{w} = 1$, which implies $\nu + \sum_{x \in \mathcal{D}} \nu_x = 1$.

Technology evolves according to

$$\ln A_{x,t} = \mathcal{S}(t; \pi_x, h_x, \kappa_x) = \pi_x \frac{1 + h_x \cdot (t_f - t_0)^{-\kappa_x}}{1 + h_x \cdot (t - t_0)^{-\kappa_x}} \tag{12}$$

where $\mathcal{S}$ is a parametric S-curve. This specification implies that $A_{x,t}$ starts at 1 and converges to $\exp(\pi_x)$—the long-run level of cost-saving gains due to the technology—at time t_f. The parameters h_x and κ_x govern the shape of the S-curve: a higher κ_x implies an S-curve with a steeper inflection; a higher h_x implies a faster adjustment. Building on the findings in Cortes (2016) summarized below, we assume that the automation of routine jobs starts in $t_0 = 1986$. Acemoglu and Restrepo (2020) estimate cost-saving gains of 30% for the automation of production jobs via industrial robots. We set $\pi_x = 30\%$ and $t_f = 2007$, which assumes similar cost-saving gains of automating other routine occupations by 2007.

We calibrate $s_{o(x)}$ (or equivalently $\nu_x = \Omega_{o(x)} \cdot s_{o(x)}$), $A_{x,t}$, and $\alpha_{0,x} = \alpha_0$ jointly to match the occupational wage decline $\theta_{o(x),t}$ from $t = 1986$ onward and the incidence of the shock by 2007. These parameters are jointly calibrated, but their values are tightly linked to the targeted moments:

- The choice of $\pi_x = 30\%$ pins the level of A_{x,t_f}.
- The occupation-level wage decline by 2007, $\theta_{o(x),t_f}$, pins $s_{o(x)}$—the exposure of each occupation to technological progress. In particular, $s_{o(x)} \to 1$ implies $\theta_{o(x),t_f} \to \ln(w_{x,t_f}/w_{t_f})$, where $w_{x,t_f}/w_{t_f}$ depends only on π_x, while $s_{o(x)} \to 0$ implies $\theta_{o(x),t_f} \to 0$.
- The time path for occupational wages $\theta_{o(x),t}$ pins the yearly time path for $A_{x,t}$ between $t_0 = 1986$ and $t_f = 2007$.
- The incidence of the shock pins α_0. The higher the incidence, the lower the reallocation rate. We calibrate a common $\alpha_0 = 3\%$ per year that matches the average incidence of 71% across occupations.[23]

[23] An alternative strategy involves calibrating a different α_x for each occupation. We do not pursue this

This procedure yields estimates for $\hat{A}_{x,t}$ for all years in the PSID since 1986. Using these values, we fit the S-curve in equation (12) via non-linear least squares.

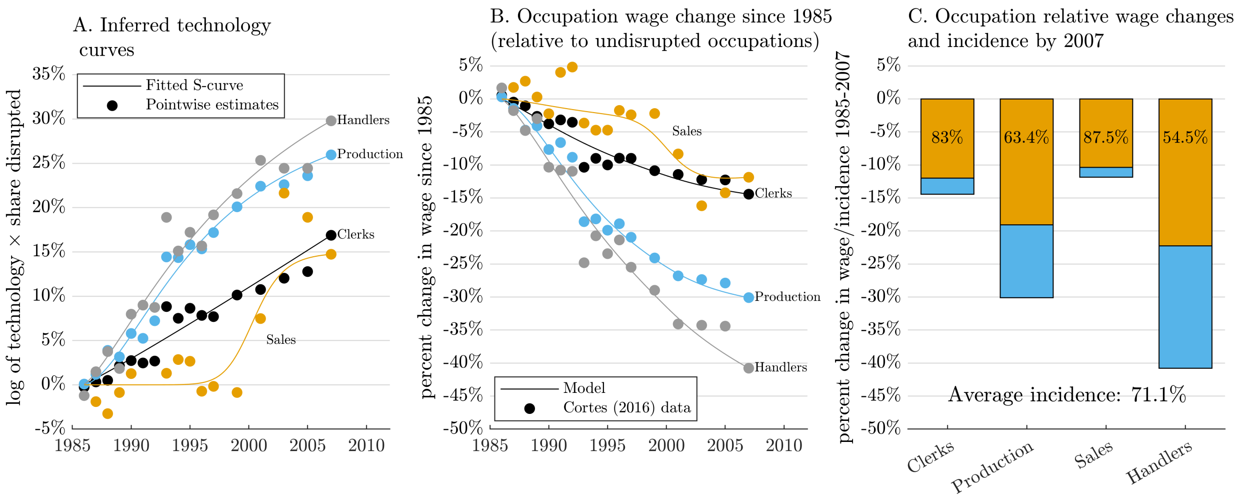

FIGURE 2: CALIBRATED PATHS FOR TECHNOLOGIES REPLACING FOR ROUTINE JOBS. Panel A reports yearly estimates of $A_{x,t}$ and their corresponding S-curve separately for each routine occupation. Panel B reports the model-implied occupational wage decline since 1985 and compares this to the estimates of $\theta_{o(x),t}$ by occupational group. Panel C reports estimates of the implied incidence of these shocks on workers that held routine occupations in 1985.

Panel A in Figure 2 reports the yearly estimates of $\hat{A}_{x,t}$ and the fitted S−curves for the disrupted segments of routine occupations. To facilitate the interpretation, these estimates are scaled by $s_{o(x)}$, so that the figure is informative of the scope and the timing of the shock across routine occupations. Panels B and C show that our model matches the empirical evidence in Cortes (2016). Panel B reports the model-implied decline in relative wages by occupation since 1986 and compares it to the empirical estimates for $\theta_{o(x),t}$ from Cortes (2016). Panel C reports incidence by occupation and shows that our model produces an average incidence of 71% as wanted.

Table 1 reports the remaining parameters. We let $u(c) = c^{1-\gamma}/(1-\gamma)$, set the inverse elasticity of intertemporal substitution $\gamma = 2$, and set the discount and interest rate to 5% per year. For the versions of our model in which households are not hand-to-mouth, we assume zero initial assets for disrupted households. This aligns with the evidence in Kaplan et al. (2014) which points to median liquid wealth for US households of $1,714 in 2010.

strategy since the occupational-specific estimates of incidence in Figure 2 for sales and transportation & material handling occupations are noisy. Instead, the average incidence of 71% is precisely estimated.

3.2 Optimal policy and the gains from gradualism

Using the formula in Proposition 3, we compute the optimal path for taxes on automation technologies since 1986. We do so for the four scenarios described in section 2: i. hand-to-mouth; ii. no borrowing and no transition risk; iii. borrowing but transition risk; iv. ex-post complete markets. These four scenarios determine the mapping between the observed decline in income documented by Cortes (2016) and matched by our model, and the unobserved marginal utility of consumption of different households over time. We report optimal taxes obtained for an utilitarian welfare function, which implies $g_x = g = 1$.[24]

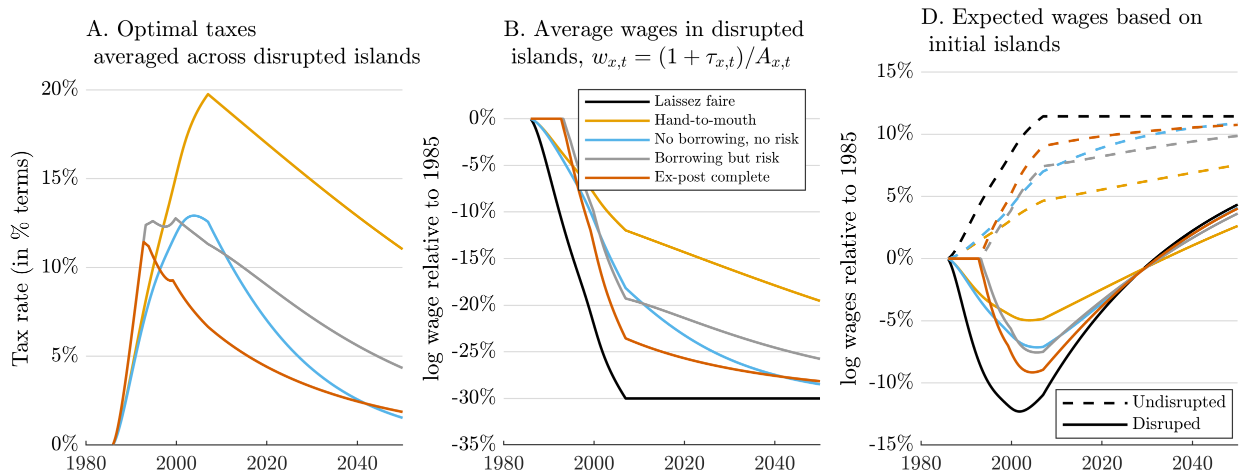

FIGURE 3: OPTIMAL TAXES AND PATH FOR WAGES ASSOCIATED WITH THE DECLINE OF ROUTINE JOBS. Panel A reports optimal taxes obtained under the four scenarios introduced in Section 1. Panel B reports the average wage in disrupted islands relative to its 1985 level. This panel also includes the status quo path where automation goes untaxed for comparison. Panel C reports the expected wage of workers initially employed in disrupted islands (solid lines) and undisrupted islands (dashed lines) relative to their 1985 levels.

Panel A in Figure 3 plots the implied optimal tax paths for these four scenarios. The orange line provides the most conservative scenario, obtained when households can borrow and insure against the risk of transitioning late. Optimal policy calls for a short-run increase in taxes on automation technologies of 11% which is then phased out, reaching a level of 4% by 2020.

Panel B plots the average wage in disrupted islands and compares it to a *Laissez Faire* scenario with no taxes. Optimal policy fully shields workers from automation technologies for 10 years, which allows them to build their savings. It then induces a more gradual

[24]This choice implies a degree of inequality aversion of 2, which is equal to the inverse of the IES. Using a strictly concave welfare function introduces a greater degree of inequality aversion and results in higher taxes.

reduction in wages at affected islands. By 2020, wages in disrupted islands are 4% higher than in the status quo.

Panel C plots expected wage paths for workers who were initially in disrupted islands. Relative to the previous panel, this one accounts for the role of reallocation. The solid black line shows a large income decline of 12% from 1985 to 2000, which aligns with the high estimated incidence in Cortes (2016). The optimal policy induces a more shallow and less persistent income drop of 7% for households in disrupted islands over 1985–2020. This panel also plots income for households in the undisrupted island. With no taxes on automation technologies, wages for unaffected workers grow gradually by 11%. The optimal policy reduces wage growth for unaffected islands and makes it more gradual.

The comparisons across scenarios illustrate two intuitive points: Optimal policy calls for more aggressive and lasting distortions when households are hand-to-mouth than when they can borrow, which aligns with the findings in Beraja and Zorzi (2022); and optimal policy calls for more aggressive and lasting distortions when households face the risk of not being able to reallocate soon enough. This is because lasting taxes insure workers against the risk of transitioning late.

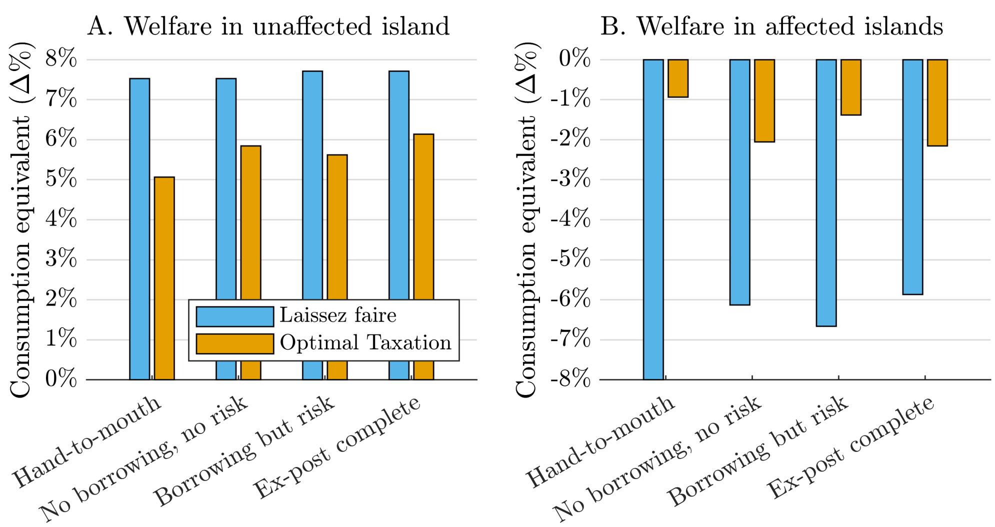

Figure 4: Welfare changes in consumption-equivalent terms, decline in routine jobs. Panel A reports welfare changes in consumption-equivalent terms for undisrupted islands under the status quo of no taxes and under the optimal policy. Panel B reports welfare changes in consumption-equivalent terms averaged across disrupted islands.

Figure 4 turns to welfare implications. Panel A reports the change in welfare in consumption-equivalent terms for households initially located in unaffected islands. Panel B reports the average change in welfare for households from disrupted islands. With no taxes, the automation of routine jobs leads to a welfare gain of 7.5% for workers who are not adversely affected and a 6–8% welfare drop for workers disrupted by this technological change. In all scenarios, the optimal policy leads to a sizable improvement in welfare for disrupted workers of 4–7 pp; while the cost for undisrupted workers is small (1–2 pp).

3.3 Endogenous effort and assistance programs

Figure 5 plots optimal taxes when reallocation effort is endogenous. Each panel considers one of our scenarios for households. For comparison, we plot the optimal tax with exogenous effort in each scenario. We then consider a simple specification of the effort elasticities where $\varepsilon_{x'',x} = 0$ for $x'' \neq x$ and $\varepsilon_{x,x} = \varepsilon > 0$ is calibrated so that moving to the optimal plan with exogenous effort would result in a reduction in effort of 10%, 20%, and 50%, respectively. These scenarios tell us how costly it is to move to the (previous) optimal policy in terms of reducing reallocation effort. The figure then provides the optimal tax in each case once we account for the induced reduction in reallocation effort. As anticipated earlier, endogenous effort leads to a more rapid phase out of taxes, and in some cases, to subsidies to the new technology in the medium run to provide incentives for reallocation.

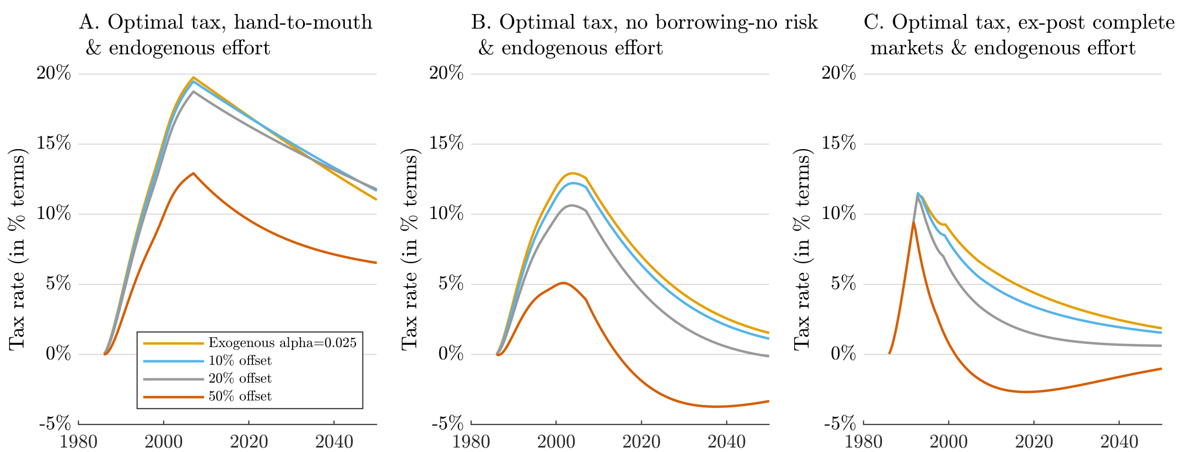

Figure 5: Optimal tax on automation technologies when reallocation effort is endogenous. Panel A considers hand-to-mouth households. Panel B considers households that share reallocation risk but cannot borrow or save outside their islands. Panel C considers the case of ex-post complete markets.

Figure 5 plots optimal taxes when work effort is endogenous and the government can create temporary assistance programs with a replacement rate $\mathcal{R}_t$. The panels consider the same scenarios for households used above but focus on the case with exogenous reallocation effort. The yellow line plots the optimal tax in the baseline scenario with no choice of work effort. The blue line plots the optimal tax when work effort is endogenous and responds to wages with an elasticity $\varepsilon_\ell = 0.5$, which matches estimates of the Hicksian elasticity of labor supply (in hours) in Chetty et al. (2011), but the government does not set any assistance program. This differs from the baseline because of endogenous changes in hours across islands.

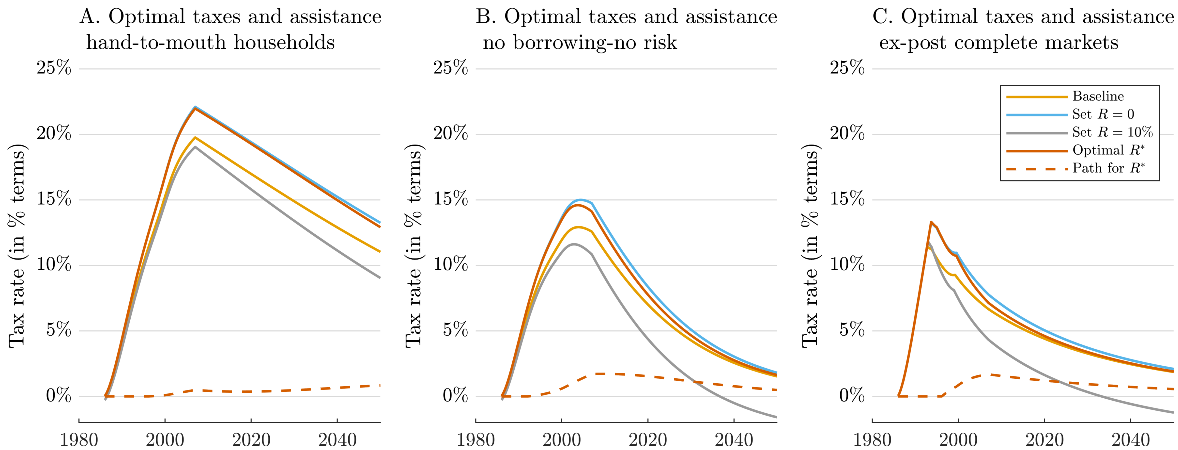

FIGURE 6: OPTIMAL TAX ON AUTOMATION TECHNOLOGIES AND ASSISTANCE PROGRAMS. Panel A considers hand-to-mouth households. Panel B considers households that share reallocation risk but cannot borrow or save outside their islands. Panel C considers the case of ex-post complete markets.

The solid and dashed orange lines plot the optimal tax on automation technologies and the optimal replacement rate of assistance programs obtained using the formulas in proposition 5. Optimal policy calls for a tepid assistance program with replacement rates of close to 2.5%. As a result, the optimal tax on automation technologies is almost unaffected. The reason why assistance programs are not used more intensively as part of the optimal policy bundle is that they are a blunt tool. Technology (and trade) generate a small set of losers. Because assistance programs cannot be targeted to these few households, they will be exploited by the large majority of workers who are not affected by technological change, generating large and costly reductions in work effort.

One potential issue with this conclusion is that, perhaps in practice, there is some imperfect level of targeting that the government can use when designing assistance programs. For example, available estimates point to replacement rates of close to 10% for workers displaced by trade.[25] These higher replacement rates might be justified if the government has access to better targeting tools. To address this issue, we also provide estimates (in gray) for the optimal tax on automation technologies obtained by fixing the replacement

[25]For example, for workers affected by Chinese import competition, Autor et al. (2014) estimate an increase in Social Security transfers that replaces 5% of their total income lost. The estimates in Table 8 of Autor et al. (2013) show that other programs, such as the federal Trade Adjustment Assistance program, unemployment benefits, federal income assistance, and education and training assistance programs, could double the replacement of lost income to 10%.

rate of assistance programs to $\mathcal{R}_t = 10\%$. These more generous assistance programs lead to a quicker phase-out of taxes on automation technologies, but do not appreciably change our conclusions.

3.4 The gains from technological gradualism

Our formula for the gains from technological gradualism implies that these are negative in the case of automation technologies, even in the absence of taxes or assistance programs. The implication is that, from a welfare point of view, advances in automation technologies since 1986 took place too slowly, and society would have benefited from more rapid advances. In sum, our exercise implies that, form a welfare point of view, the somewhat rapid advances in automation technologies seen since 1986 are a welcomed development. A short run tax on automation technologies would have made things better by inducing a more gradual wage path for exposed workers and raising additional tax revenue. According to our calibration, there is no point in smashing the machines; one can just tax them.

4 Application II: The China Shock

4.1 Description, Empirical Evidence, and Calibration

Description: there are 21 islands. Islands 2–21 are in $\mathcal{D}$ and represent segments of 2-digit manufacturing industries $i(x)$ disrupted by import competition from China. As before, we assume that only a fraction $s_{i(x)}$ of the products or varieties produced by industry $i(x)$ are exposed to Chinese competition. This implies that the disrupted island x associated with industry $i(x)$ accounts for a fraction $\nu_x = s_{i(x)} \cdot \Omega_{i(x)}$ of initial value added, where $\Omega_{i(x)}$ is the value added share of that industry.[26] We calibrate $s_{i(x)}$ to match the scope of the disruption brought by the China Shock in each industry. Table 2 lists all 2-digit manufacturing industries, their SIC codes, and their shares of value added in 1991. The first island represents segments of manufacturing industries that were not exposed to Chinese

[26] As before, one may consider a mass 1 of islands partitioned into industries. Island x belong to industry $i(x)$, and there is a mass $\Omega_{i(x)}$ of islands in each industry. One can think of each island as producing one differentiated variety. We can then model the rise of Chinese imports as resulting from improvements in the productivity of China at a share $s_{i(x)}$ of the islands associated with industry $i(x)$. The remaining share of varieties are shielded from Chinese competition. For example, Holmes and Stevens (2014) show that import competition affected primarily large firms engaged in the production of standardized goods within exposed industries. Those firms make up the disrupted islands. This aligns with the fact that there are sizable differences in Chinese import penetration across detailed industries and products, even within the 2-digit manufacturing industries used in our analysis (see Autor et al., 2013).

competition plus all non-manufacturing industries.

The empirical evidence in Autor et al. (2013) and Autor et al. (2014): These papers provide two key moments. Autor et al. (2013) measure Chinese import penetration by industry using data from Comtrade for 1991 to 2007. They document that the increase in Chinese imports within industries is highly correlated across advanced countries, which suggest that the China Shock is driven by improvements in Chinese exports' productivity.

Panel A in Figure 7 summarizes their data and plots the *change in normalized import shares* for manufacturing industries. This is computed as

$$\text{Change in normalized import share}_i = \frac{1}{\Omega_i} \cdot (m_{i,t}/y_t - m_{i,t_0}/y_{t_0}),$$

where $m_{i,t}$ denotes the value of Chinese imports in industry i. Normalizing by Ω_i makes these estimates comparable across industries. Normalizing imports by GDP at time t ensures that this measure does not capture a mechanical increase in imports driven by US economic growth. While Chinese imports started to increase for some industries since the early 90s, there is a more pronounced and pervasive increase in 2000–2007.[27]

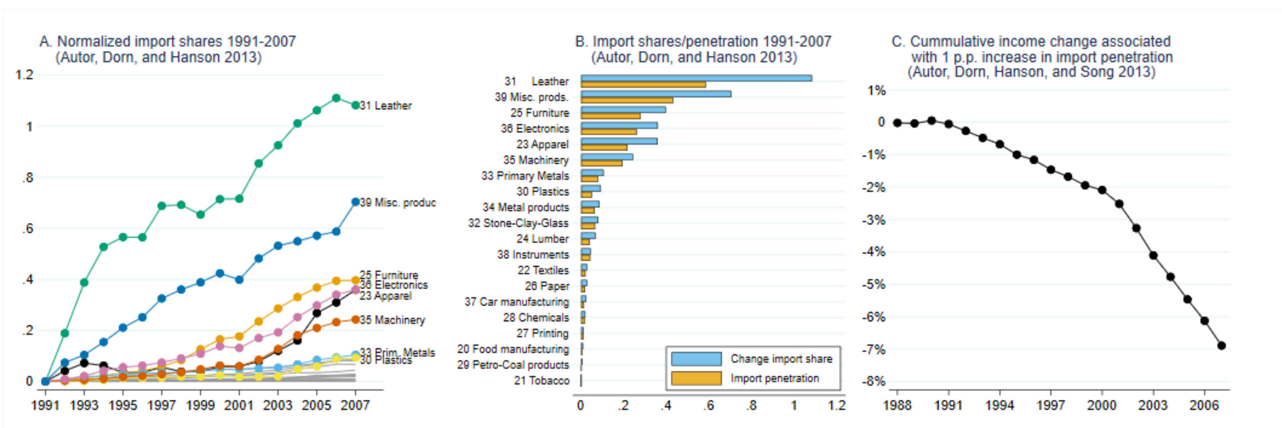

FIGURE 7: MEASURES OF IMPORT COMPETITION FROM CHINA AND THE EFFECT OF THE CHINA SHOCK ON WAGE GROWTH OF EXPOSED WORKERS. Panel A reports estimates of normalized import shares over 1991–2007 using the data from Autor et al. (2013). Panel B reports the increase over 1991–2007 in normalized import shares and compares this to Autor et al. measure of import penetration. Panel C reports estimates from Autor et al. (2014) of the effects of a 1 pp increase in import penetration on cummulative future income growth of exposed workers.

Panel B lists the increase in normalized import shares over 1991–2007 for 2-digit man-

[27] Part of the acceleration can be attributed to China's accession to the World Trade Organization (WTO) at the end of 2001, which resulted on the US granting Permanent Normal Trade Relations (PNTR) to China (see Pierce and Schott, 2016). This created an incentive for US firms to open plants abroad, which maps to a rise in $A_{x,t}$ in our model.

ufacturing industries. On average, Chinese imports rose by 11 pp of manufacturing value added, though there is sizable heterogeneity, with industries such as apparel, leather products, and miscellaneous manufacturing products (i.e., toys, games, and jewelry) experiencing 40–110 pp increases in import competition. The increase in import shares is informative of the share of segments disrupted $s_{i(x)}$ and the path of Chinese productivity $A_{x,t}$ in these islands.

Autor et al. (2014) provide evidence that workers initially employed in industries facing more import competition from China experienced lower income growth after 1991. They use data from the Social Security Administration to estimate the model:

$$\text{Cumulative earnings}_{i,h,t} = \beta_{\text{w},t} \cdot \Delta\text{IP } 91\text{–}07_i + \theta \cdot \text{X}_{h,i} + u_{h,i}, \tag{13}$$

which explains cumulative earnings (relative to baseline earnings) for person h employed in industry i since 1991 and up to time t as a function of import penetration in their industry, covariates, and an error term.[28] Cumulative earnings relative to baseline income are measured as

$$\text{Cumulative earnings}_{t,h,i} = \frac{\sum_{t'=1992}^{t} \text{Labor income}_{h,i,t'}}{\text{Labor income}_{h,i,t_0}} \text{ where } t_0 = 1991.$$

Their measure of import penetration is similar to the normalized share of imports introduced above, but differs in that it normalizes import growth by the total consumption in the US of goods produced by industry i in 1991 and does not normalize imports by total GDP in the US over time. For comparison, Panel B also plots their import penetration measure. Both measures are highly correlated (correlation coefficient=0.99), but normalized import shares are convenient to work with in our model.

Panel C in Figure 7 plots the estimates of $\beta_{w,t}$ in equation (13), obtained from Figure III in Autor et al. (2014). The estimates show no pre-trends in labor income prior to 1991. Labor income then declines in relative terms, and by 2007, workers who were employed in industries with a 7.5 pp higher import penetration (the average level in manufacturing industries) saw their cumulative income from 1992 to 2007 decline by 50% of their baseline annual income (7.5× 6.8) relative to workers not exposed to Chinese competition. This sizable effects suggests that the China Shock had considerable incidence on workers

[28] Autor et al. (2014) report IV estimates of equation (13). They instrument Chinese import penetration in the US using Chinese import penetration in the same industry but in other high-income countries. This strategy isolates the variation in Chinese imports coming from changes in supply. Their first-stage results suggest that 80% of variation in import penetration can be attributed to supply-side forces such as rising manufacturing productivity in China.

employed at disrupted industries, which points to limited opportunities for workers to reallocate. Autor et al. (2014) also show that, even among workers employed in the same 2-digit industry, all of the incidence of the China Shock falls on workers that specialized in the detailed industries experiencing the biggest increase in import penetration, which points to limited opportunities for reallocation even within an industry.[29]

Calibration: Aggregate output y_t is given by the CES in (11). We take a value of $\sigma = 2$, which aligns with the median elasticity of substitution across varieties at the level of aggregation used in our analysis (see, for example, Broda and Weinstein, 2006).[30] As before, we normalize $\bar{w} = 1$.

Technology evolves according to the S-curve in (12). Building on the findings in Autor et al. (2013) and Autor et al. (2014), we assume that the China Shock starts in $t_0 = 1991$.[31] We also set a common value of $\pi_x = \pi$ across industries representing the cost-saving gains from trading with China by $t_f = 2007$. We calibrate π to match empirical estimates of price declines generated by Chinese import competition in US markets. Bai and Stumpner (2019) document that a 1% decrease in the share of goods produced domestically in an industry is associated with a 0.36–0.5% decline in consumer prices. In our model, the relationship between industry prices and domestic production shares satisfies

$$\Delta \ln P_i \approx \text{constant} + \pi \cdot \Delta \ln \text{share domestic production}_i.$$

Intuitively, substituting a variety produced domestically for a Chinese variety generates a cost-saving gain of π per variety substituted. We set $\pi = 50\%$, to match the upper end of the estimates in Table 1 of Bai and Stumpner (2019).[32]

[29] This is precisely what our model generates. In our model, all of the incidence of the China Shock falls on workers in disrupted islands, and not on other workers in their same industry. This is different from what one would get in a model where labor is freely mobile across jobs within an industry. In these alternative models, the welfare losses would be smaller but shared among many more workers.

[30] Recall that disrupted islands correspond to specific products or varieties within each 2-digit SIC industry. The closest level of aggregation considered in Broda and Weinstein (2006) is the elasticity of substitution among products within a 3-digit SIC level. It is also important to note that $\sigma - 1$ is not equivalent to the *trade elasticity* that features in various trade models. In our model, the equivalent of a trade elasticity for island x is trade elasticity $= \sigma \cdot \frac{y_{x,t}}{k_{x,t}} - 1 > \sigma - 1$. This elasticity exceeds $\sigma - 1$, because our model with islands getting fully out-competed features an extensive margin of trade (infinite elasticity) and an intensive margin (elasticity $\sigma - 1$). Our calibration produces an *average* trade elasticity for disrupted islands of 4.8.

[31] There was a small level of pre-existing trade with China before 1991. Appendix A.5 explains how we extend our model to capture pre-existing trade and how we deal with it in our calibration.

[32] Appendix A.5 derives this approximation. Another way of understanding the choice of π is by noting that this equals the lower Chinese cost at varieties supplied by China in the US. In a model where firms differ in their productivity and their productivity has a Pareto distribution (e.g., Chaney, 2008), this

We calibrate $s_{o(x)}$ (or equivalently $\nu_x = \Omega_{o(x)} \cdot s_{o(x)}$), $A_{x,t}$, and $\alpha_{0,x} = \alpha_0$ jointly to match the increase in the share of Chinese imports over 1991–2007 and the incidence of this shock on workers in disrupted industries. These parameters are jointly calibrated, but their values are tightly linked to the targeted moments:

- The choice of $\pi_x = 50\%$ pins the level of A_{x,t_f}.
- The level of normalized import shares by 2007 pins $s_{i(x)}$—the exposure of each industry to Chinese import competition. In particular, normalized import shares by 2007 are proportional to $s_{i(x)}$, so that industries with the greatest increase in import competition have a higher share of products and segments exposed.
- The time path for normalized imports pins the path for $A_{x,t}$ between $t_0 = 1991$ and $t_f = 2007$.
- The incidence of the shock pins α_0. The higher the incidence, the lower the reallocation rate. We calibrate a common $\alpha_0 = 1.8\%$ per year that matches the estimates of cumulative wage growth from Autor et al. (2014) by 2007 for affected workers.

This procedure yields estimates for $\hat{A}_{x,t}$ for all years since 1991. Using these values, we fit the S-curve in equation (12) via non-linear least squares.

Panel A in Figure 8 reports yearly estimates of $\hat{A}_{x,t}$ and the fitted S−curves for the disrupted segments of manufacturing industries. To facilitate the interpretation, these estimates are scaled by $s_{i(x)}$, so that the figure is informative of the scope and the timing of the shock across industries. We report these for the top 8 industries with highest exposure to import competition. Panels B and C show that our model matches the empirical evidence in Autor et al. (2013) and Autor et al. (2014). Panel B reports the model-implied rise in imports since 1991 and compares it to the estimates from Autor et al. (2013). Panel C reports the incidence of the shock on workers associated with a 1 pp increase in Chinese import penetration. Our model matches the cumulative income decline for exposed workers by 2007. Though not targeted, our model also matches the time path of income for exposed workers during the intervening years.

average cost advantage equals $\exp(\pi) = \frac{\zeta}{\zeta-1}$, where ζ is the tail index of the Pareto distribution. Recall that in these models, $1/\zeta$ also gives the elasticity between prices and domestic shares. This implies that $\ln \pi \approx 1/\zeta \approx 36\% - -50\%$ for the estimates in Bai and Stumpner (2019). The trade literature provides additional estimates for ζ exploiting variation in tariffs across countries, which place it around 4–5, implying $\pi = 20 - 25\%$ (see Head and Mayer, 2014, for a meta-analysis of the estimates). We prefer the estimates in Bai and Stumpner (2019) because these describe trade between the US and China.

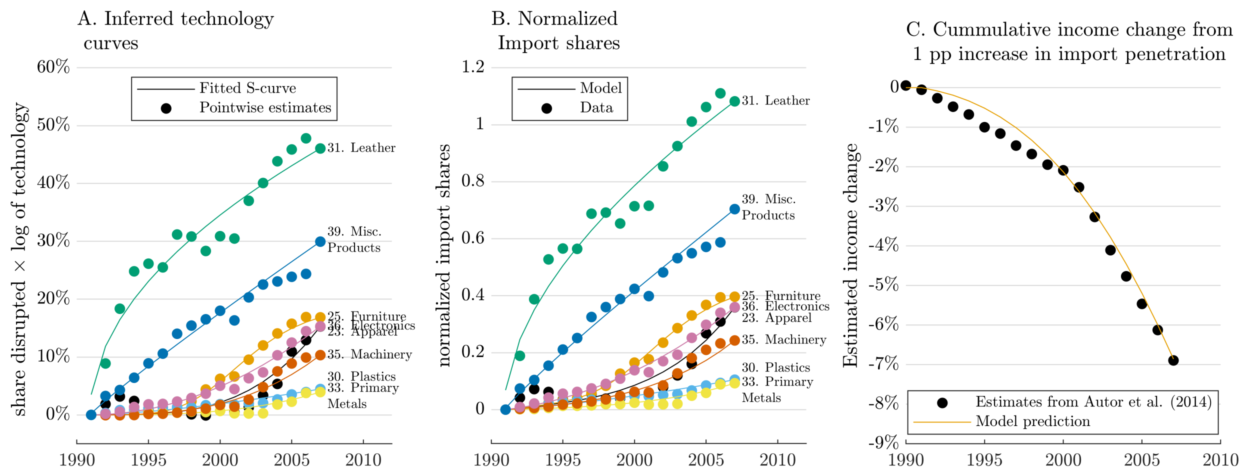

FIGURE 8: CALIBRATED PATHS FOR ADVANCES IN CHINESE PRODUCTIVITY FOR HIGHLY-EXPOSED INDUSTRIES. Panel A reports yearly estimates of $A_{x,t}$ and their corresponding S-curve separately for each routine occupation. Panel B reports the model-implied import shares since 1991 and compares this to the estimates in Autor et al. (2013). The figure provides these data for the top 8 industries with the highest exposure to Chinese competition. Panel C reports estimates of the implied incidence of these shocks on workers in industries with a 1 pp higher exposure to import penetration and compares it to the estimates in Autor et al. (2014).

Table 2 reports the remaining parameters used in our calibration, which are the same used in the application of our model to the decline in routine jobs.

The calibration for the decline in routine jobs and the China Shock vary in details but exploit similar information. In both cases, the rate of reallocation determines the incidence of these shocks, which we calibrate to match the empirical estimates of future wage growth for workers who were initially at disrupted industries or occupations. Through the lens of our model, the high incidence of these shocks on workers points to limited opportunities to reallocate. We then show that one can use trends in occupational wages or imports by industry to recover the time path of $A_{x,t}$.

4.2 Optimal policy and the gains from gradualism

Using the formula in Proposition 3, we compute the optimal path for taxes on Chinese imports starting in 1991. We do so for the same four scenarios for households introduced before. These four scenarios determine the mapping between the observed decline in income documented by Autor et al. (2014) and matched by our model, and the unobserved marginal utility of consumption of disrupted households over time. As before, we report results for a utilitarian welfare function.

Panel A in Figure 3 plots the implied optimal tax paths. The orange line provides the

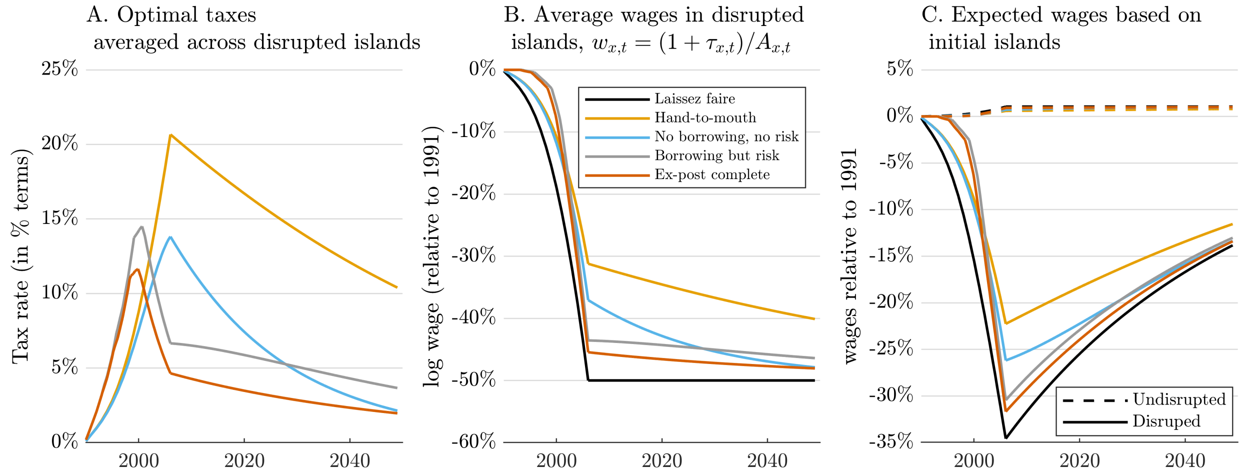

Figure 9: Optimal taxes and path for wages associated with the China Shock. Panel A reports optimal taxes obtained under the four scenarios introduced in Section 1. Panel B reports the average wage in disrupted islands relative to its 1991 level. This panel also includes the status quo path where Chinese imports go untaxed for comparison. Panel C reports the expected wage of workers initially employed in disrupted islands (solid lines) and undisrupted islands (dashed lines) relative to their 1991 levels.

most conservative scenario, obtained when households can borrow and insure against the risk of transitioning late. Optimal policy calls for a short-run increase in taxes on Chinese imports of 12% which is then phased out over time, and reaches a level of 4% by 2020. Panel B plots the resulting wages in disrupted islands and compares it to their level in a Laissez-faire world. Optimal policy fully shields workers from trade competition for 5 years, which allows them to build their savings. It then induces a more gradual reduction in wages at affected islands. By 2020, wages in disrupted islands are 5% higher than in a world with no taxes on Chinese imports.

Panel C plots expected wage paths for workers who were initially in disrupted islands. Relative to the previous figure, this one accounts for the role of reallocation. The solid black line shows a large income decline of 30% from 1985 to 2000, which aligns with the high estimated incidence in Autor et al. (2014). The optimal policy induces a more modest income drop for households in disrupted islands over 1985–2040. This panel also plots income for households in the undisrupted island over time. With no taxes on Chinese imports, wages for unaffected workers grow gradually by 1% thanks to the gains from trade.

Figure 10 turns to welfare. Panel A reports the change in welfare in consumption-equivalent terms for households that were initially in unaffected islands. Panel B reports the average change in welfare for households from disrupted islands. With no taxes on

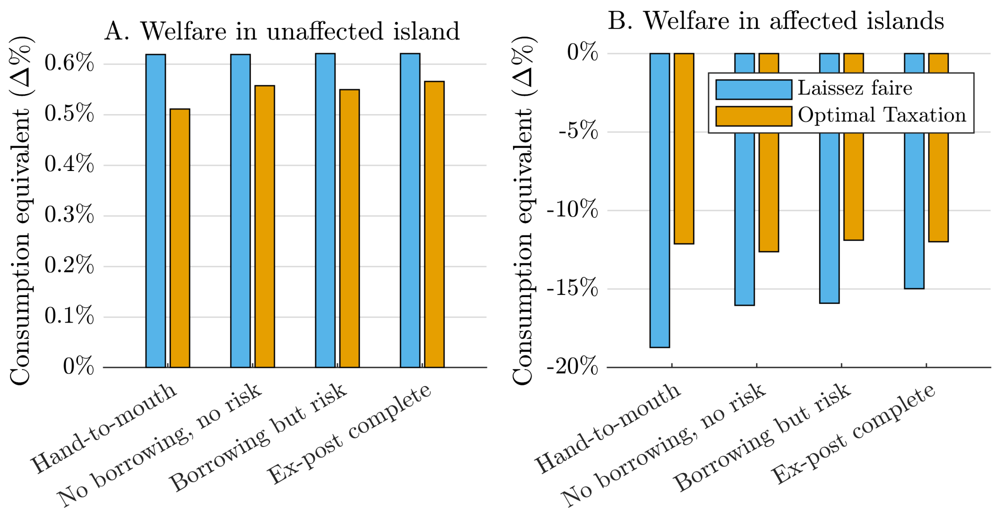

FIGURE 10: WELFARE CHANGES IN CONSUMPTION-EQUIVALENT TERMS, THE CHINA SHOCK. Panel A reports welfare changes in consumption-equivalent terms for undisrupted islands under the status quo of no taxes and under the optimal policy. Panel B reports welfare changes in consumption-equivalent terms averaged across disrupted islands.

imports, the China Shock leads to welfare gains of 0.6% for workers who are not exposed to international competition and a 15% welfare drop for those exposed to it (though these workers represent only 1.6% of the US workforce). The small gains from trade align with the trade literature and reflect the (still) low aggregate levels of Chinese import penetration in the US. For example, Galle et al. (2022) estimate gains from trade of 0.3% from trade with China. In all scenarios, the optimal policy leads to a sizable improvement in welfare for disrupted workers of 3–6 pp; while the cost for undisrupted workers of slowing down the China Shock is small (0.05–0.1 pp).

4.3 Endogenous effort and assistance programs

Figure 11 plots optimal taxes when reallocation effort is endogenous. Each panel considers one of our scenarios for households. For comparison, we plot the optimal tax with exogenous effort in each scenario. Endogenous effort leads to a quicker phase out of taxes, and in some cases, to import subsidies by 2040 to provide incentives for reallocation.

Figure 11 plots optimal taxes when work effort is endogenous and the government can create temporary assistance programs with a replacement rate $\mathcal{R}_t$. The panels consider the same scenarios for households used above but focus on the case with exogenous reallocation effort. The yellow line plots the optimal tax in the baseline scenario with no choice of work effort. The blue line plots the optimal tax when work effort is endogenous and responds to wages with an elasticity $\varepsilon_\ell = 0.5$.

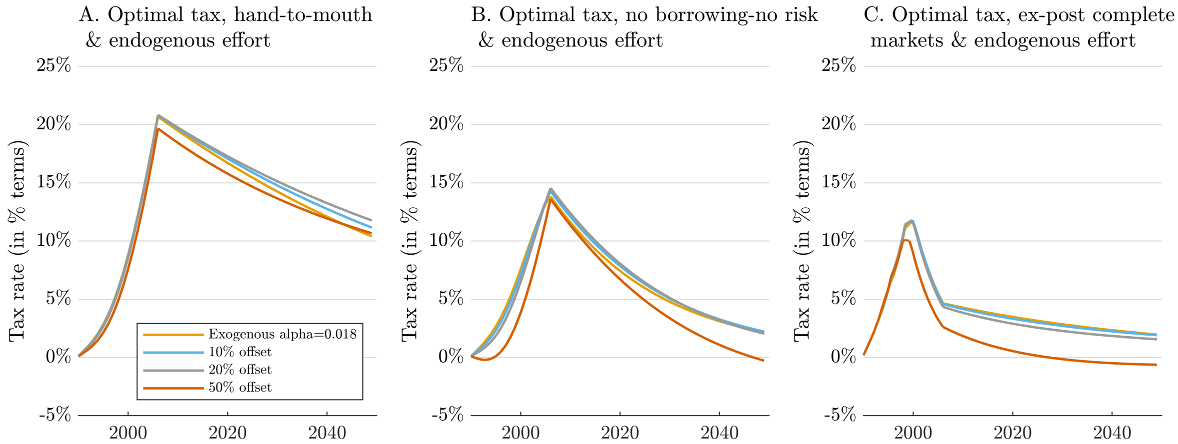

FIGURE 11: OPTIMAL TAX ON CHINESE IMPORTS WHEN REALLOCATION EFFORT IS ENDOGENOUS. Panel A considers hand-to-mouth households. Panel B considers households that share reallocation risk but cannot borrow or save outside their islands. Panel C considers the case of ex-post complete markets.

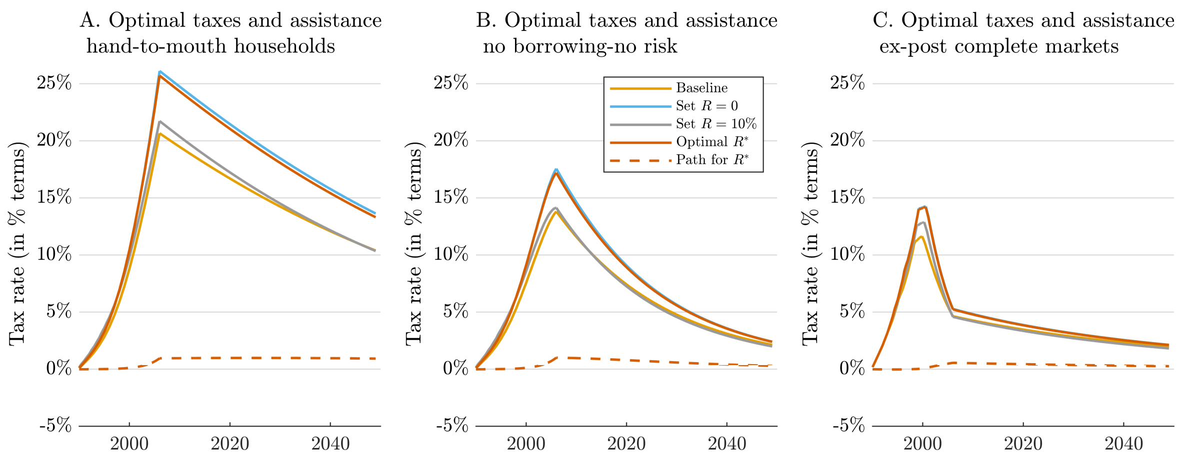

FIGURE 12: OPTIMAL TAX ON CHINESE IMPORTS AND ASSISTANCE PROGRAMS. Panel A considers hand-to-mouth households. Panel B considers households that share reallocation risk but cannot borrow or save outside their islands. Panel C considers the case of ex-post complete markets.

The solid and dashed orange lines plot the optimal tax on automation technologies and the optimal replacement rate of assistance programs obtained using the formulas in proposition 5. Optimal policy calls for a tepid assistance program with replacement rates of close to 2%. As a result, the optimal tax on Chinese imports is almost unaffected by the availability of assistance programs. As before, these programs are too blunt to deal with the fact that trade competition generates a small set of losers. As before, we also provide estimates (in gray) for the optimal tax on automation technologies obtained by fixing the replacement rate of assistance programs to $\mathcal{R}_t = 10\%$. These more generous assistance programs lead to a quicker phase-out of taxes on automation technologies, but do not appreciably change our conclusions.

4.4 The gains from technological gradualism

Our formula for the gains from technological gradualism implies that these are negative in the case of the China Shock, even in the absence of taxes or assistance programs. The implication is that, from a welfare point of view, advances in Chinese import penetration since 1986 took place too slowly, and society would have benefited from more rapid advances. In sum, our exercise implies that, form a welfare point of view, the rapid development of the Chinese exporting sector was a welcomed development. A short run tax on imports would have made things better by inducing a more gradual wage path for exposed workers and raising additional tax revenue.

5 Application III: Trade Liberalization in Colombia

5.1 Description, Empirical Evidence, and Calibration

Our final application explores Colombia's trade liberalization in 1990. Before the reform, Colombia had arresting levels of trade protection, with nominal tariffs on manufacturing imports of 40%, and effective tariffs—which account for other barriers and surcharges—reaching levels of 75% (see Goldberg and Pavcnik, 2005; Eslava et al., 2013). Pre-reform levels of trade protection also featured vast dispersion both within and across industries, with apparel and shoes enjoying effective tariffs of close to 120%, and intermediate-goods imports being subsidized or enjoying no protection. With the government of President César Gaviria in 1990, Colombia embarked in an ambitious program of economic reforms that included liberalizing labor markets and opening up to trade.[33] The initial plan was for trade liberalization to be implemented gradually, with the motivation that this would give workers and firms time to adjust. But concerns about the credibility of the reform process and the potential for the emergence of political roadblocks led to a swift implementation (see Edwards and Steiner, 2008).[34] By 1992, Colombia had reduced all nominal tariffs

[33]These reforms were part of a broader regional movement away from decades of protectionism under the auspices of *import substitution programs*. Chile was at the forefront of the reform movement, and had a gradual liberalization process starting in 1975. Subsequent reformers, such as Argentina, Colombia, Costa Rica, and Nicaragua embraced a more rapid reform process. For example, Nicaragua reduced nominal tariffs from 110% to 12% from 1990 to 1992. See Edwards (1994) for more on the reform movement in Latin America during this period.

[34]The first concern was that a gradual reform would not do enough to convince Colombian firms to upgrade or restructure their operations. The second concern was that a gradual reform could be stopped

to common international levels of close to 13% and removed almost all additional trade barriers, leading to a new trade structure with uniform effective taxes of 25% across most manufacturing industries, with the exception of imported food products.

Figure 13 depicts the large and rapid decline in effective tariffs starting in 1990 for 3-digit manufacturing industries in Colombia. On average, the distortions induced by effective tariffs—or $1 + \tau_{x,t}$ in our model—declined by 37.5% by 1992. The drop in protection was met by a sharp increase in imports, with import penetration as a share of GDP rising from 9% in 1989 to 14% in 1993, and settling at 16% by 2005. The middle panel in Figure 13 summarizes the total decline in effective tariffs over 1989–2002 and the resulting increase in normalized import shares over this period by industry. The right panel shows the almost immediate increase in imports as a share of GDP following the liberalization.

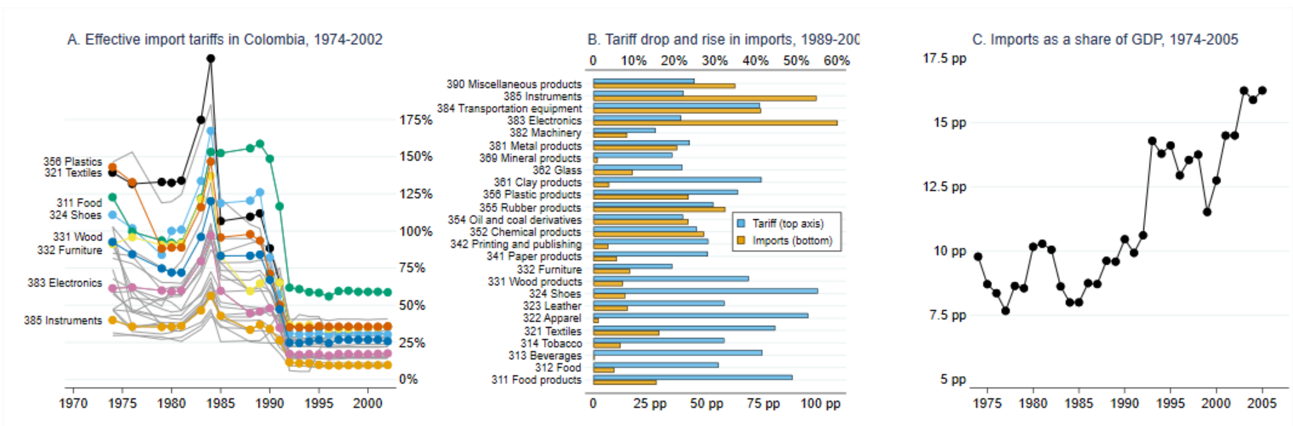

Figure 13: Trade liberalization in Colombia. Panel A reports the time series for effective tariffs in each 3-digit manufacturing industry over 1974–2002. Panel B reports the percent decline in tariffs, defined as the percent change in $1 + \tau_{x,t}$ from 1989 to 2002 and the associated increase in normalized import shares for each industry. Panel C reports imports as a share of GDP. Data on tariffs comes from Colombia's *Ministerio del Comercio* and data on imports comes from the *Departamento Nacional de Planeación*, DNP.

To map the theory to the data, we must deal with two aspects of Colombia's trade liberalization. First, tariffs were not lowered to zero—the optimal level in our model. The reform lowered tariffs to "internationally acceptable levels" of $\tau_{x,t_f} = 13\%$ for nominal taxes and $\tau_{x,t_f} = 25\%$ for effective taxes across most industries and products. Because we do not want to confound the gains from gradualism with issues related to the optimal long-run level of import tariffs, we assume that there is another distortion in the economy that makes the post 1990 level of protection optimal and report series for the net optimal tariff,

before achieving the desired level of liberalization if the political climate changed or affected groups managed to organize against the reform.

defined as $1+\tau^{net}_{x,t}=\frac{1+\tau_{x,t}}{1+\tau_{x,t_f}}$..[35]

Second, there were large trade volumes in some industries before the reform. This was particularly the case for durable manufacturing and industries producing capital goods, many of which had been liberalized prior to 1989. We account for this feature of Colombia's tariff and trade structure by assuming that some goods and services were already produced abroad by 1989 and these goods experienced no change in tariffs during Colombia's trade liberalization. Instead, Colombia's trade liberalization worked at the extensive margin: by making it profitable to import a widening range of products that used to be produced domestically.[36]

Calibration: We consider an economy with 26 islands. Islands 2–26 represent segments of manufacturing industry $i(x)$ that survived trade competition because of the protection granted by the high tariffs in 1989, but were out-competed following the trade liberalization.[37] These segments account for a share $s_{i(x)}$ of industry i. We assume that the initial level of protection in industry $i(x)$ is set at the minimum level required to ensure that imports did not disrupt island x, which implies

$$\frac{A_{x,t_0}}{1+\bar{\tau}_{x,t_0}}=\bar{w}, \qquad \text{for } t_0=1989 \text{ and } x\in\mathcal{D}.$$

This allow us to recover the pre-tax productivity of imports as $A_{x,t_0}=(1+\bar{\tau}_{x,t_0})\cdot\bar{w}$ for all disrupted islands.

[35] Formally, we assume that imported goods in island x have an implicit an exogenous subsidy of $1+\tau_{x,t_f}$, where τ_{x,t_f} denotes the effective tax rate in 2002. In this case, the optimal tax is $1+\tau_{x,t}=(1+\tau^{net}_{x,t})\cdot(1+\tau_{x,t_f})$, where $\tau^{net}_{x,t}$ is the optimal tax characterized in Proposition 3 obtained after redefining technology to $A^{net}_{x,t}=A_{x,t}/(1+\tau_{x,t_f})$.

[36] A different interpretation is that trade liberalization worked primarily at the intensive margin: by reducing tariffs for products that were already imported and not produced domestically. This would imply larger welfare gains and fewer or no distributional costs of the reform. This alternative interpretation does not provide an accurate representation of Colombia's trade liberalization for several reasons. First, Colombia's protection structure featured vast differences in tariffs across products within detailed industries, which suggest that most of the pre-existing trade took place in pockets of industries and products that benefited from low protection or subsidies in some cases. For example, 65% of the variation in tariffs in 1989 was between products in the same 3-digit manufacturing industry. Likewise, industries with the highest levels of trade by 1989 had 10–20% of products already exempted from tariffs or subsidized. Finally, there was no increase in import penetration or normalized import shares in the industries with the largest share of pre-existing trade, such as industrial chemicals and petroleum products (codes 351, 353) or primary metals (codes 371, 372). In fact, initial import shares only explain 10% of the variation in subsequent import penetration after 1989, which is the opposite of what one would have expected if trade liberalization involved mainly changes at the intensive margin.

[37] We exclude the intermediate-goods industries 351 "industrial chemicals", 353 "oil refined products", 371 "steel and iron" and 372 "primary metals" from our exercise because they were already liberalized prior to 1990 and experienced no increase in import penetration since then.

As in the China-Shock application, we set $\sigma = 2$ and calibrate the shares $s_{i(x)}$ to match the observed increase in normalized import shares by industry in response to the drop in tariffs during the trade liberalization. Conditional on the decline in effective tariffs, industries with a larger share of exposed segments, $s_{i(x)}$, should see a more pronounced increase in normalized import shares post 1990.[38]

For this application, we do not have data on the incidence of trade liberalization on workers who specialized in exposed industries. However, Goldberg and Pavcnik (2005) provide a related piece of evidence. Exploiting variation in changes in protection over time across Colombian industries, they show that a 10 percent decrease in tariffs is associated with a decline in industry wage premiums of 1%. Their estimate of the decline in the manufacturing wage premium contains information on α. Higher values of α imply that the average manufacturing worker is more likely to have reallocated to non-disrupted manufacturing jobs, which implies a less pronounced impact of trade liberalization on manufacturing wages. In the limit with $\alpha = \infty$, workers earn the same wage at all industries and trade does not affect industry wage premia. A value of $\alpha = 3\%$ matches Goldberg and Pavcnik (2005) estimates.

With this calibration at hand, we set $t_0 = 1989$—the year before Gaviria's reforms—and feed the observed path for effective tariffs to obtain the path for wages, imports, and aggregates under the reform.

5.2 Optimal trade liberalization

We use the optimal tax formula in Proposition 3 to compute a counterfactual path for effective tariffs that maximizes social welfare under different assumptions on households' saving and consumption behavior. Figure 14 reports our findings. Panel A depicts the observed path for tariffs following the 1990 trade liberalization and compares it to the optimal path implied by Proposition 3. From a welfare point of view, the reform was done too swiftly. Let's compare the observed path with the optimal path when households are hand to mouth but share the reallocation risk. The optimal trade liberalization calls for a quick drop in net tariffs to 13% and then a slow tariff decline reaching a level of 5%

[38] The assumption behind this procedure is that the observed increase in import penetration in the 90s was entirely due to the large reduction in effective tariffs. We see this as a reasonable approximation, especially when considering the vast drop in tariffs, and taking into account the fact that, as shown in Figure 13, the decline in tariffs was met by an almost immediate rise in imports. Our procedure matches exactly the rise in imports across all industries from 1989 to 2002, except for 385 (scientific and medical instruments) and 383 (electronics). For these two industries, the restriction that $s_{i(x)} \leq 1$ binds and our model understates the large increase in imports by 34 pp and 24 pp, respectively.

by 2010. Instead, the 1990 trade liberalization featured an almost-immediate drop in net tariffs to zero.

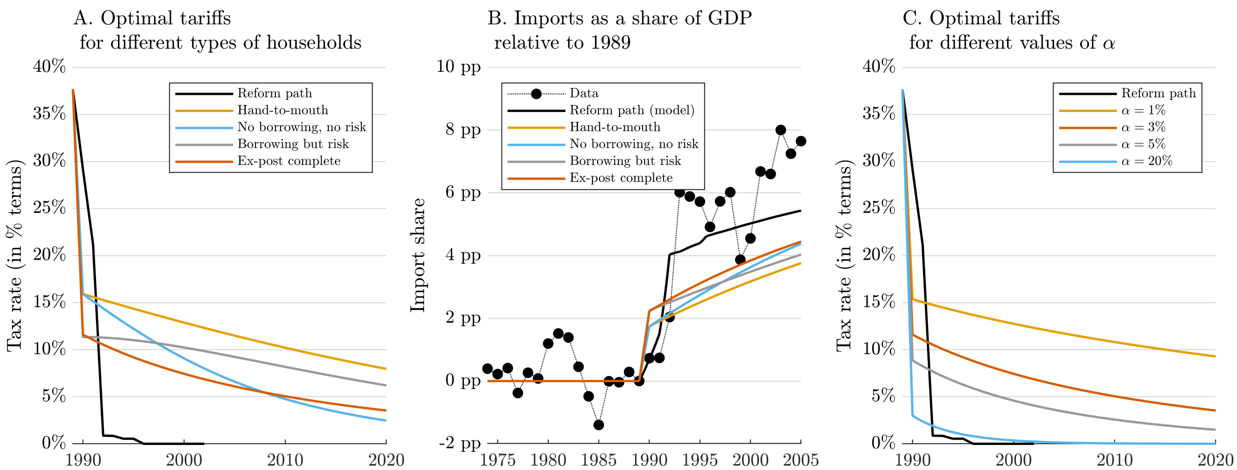

FIGURE 14: OPTIMAL TARIFFS AND OBSERVED TARIFFS FOR COLOMBIA'S TRADE LIBERALIZATION. Panel A reports optimal taxes obtained under the four scenarios introduced in Section 1. Panel B reports imports as a share of GDP relative to 1989 under the observed and the optimal policy. Panel C reports optimal taxes obtained for different values of the reallocation rate α when households face ex-post complete markets.

Panel B shows the observed path for imports and the counterfactual path under the optimal policy. We see that both in the model and data, imports rose rapidly after the 1990 trade liberalization. Instead, the optimal policy would induce a more gradual increase in imports over time in all scenarios.

Panel C reports the optimal tariff path for different values of α and under the conservative assumption that households face ex-post complete markets (they can borrow and face no transition risk). The swift reform conducted in Colombia (and in much of Latin America during that period) is only justified for high reallocation rates of 20% per year—an order of magnitude larger than our estimate.

Despite the fact that some industries enjoyed more protection than others, the optimal trade liberalization calls for a similar gradual path across sectors that retains some of the dispersion in tariffs across the transition. This is the opposite of what one would get on pure efficiency grounds, which call for a more aggressive dismantling of barriers in more heavily protected sectors (see Mussa, 1984, for a discussion of issues of tariff dispersion in trade reforms).

Figure 15 reports the welfare gains and costs from trade liberalization under different scenarios and paths for tariffs. Colombia's trade liberalization brought welfare gains of 2.2% for unaffected workers and large welfare losses of 12–14% for disrupted workers (3.4%

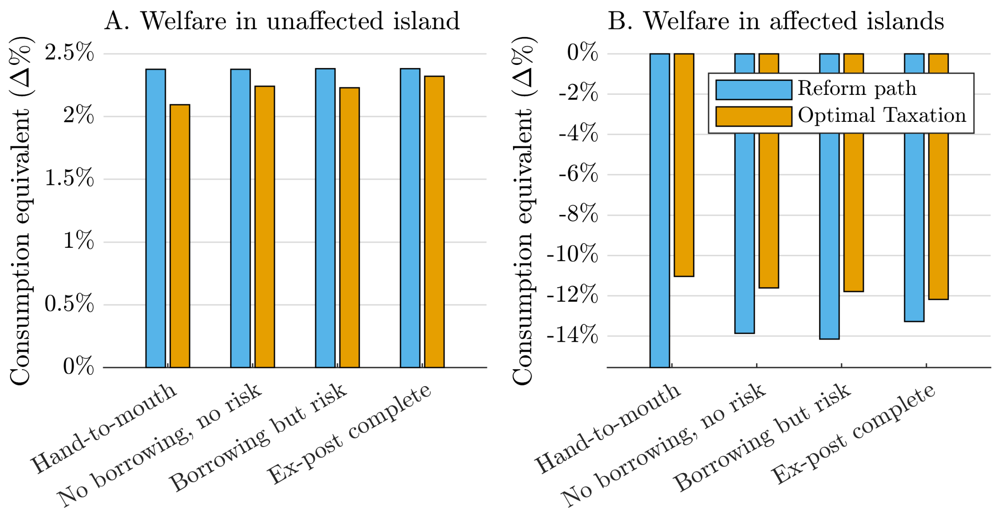

FIGURE 15: WELFARE CHANGES IN CONSUMPTION-EQUIVALENT TERMS, COLOMBIA'S TRADE LIBERALIZATION. Panel A reports welfare changes in consumption-equivalent terms for undisrupted islands under the observed policy and the optimal policy. Panel B reports welfare changes in consumption-equivalent terms averaged across disrupted islands under the observed policy and the optimal policy.

of the workforce in this case). A more gradual approach would mitigate the losses by 0.5–4 pp, depending on assumptions on households' ability to borrow and share risks, and come at a small welfare cost for unaffected workers of 0.05-0.1 pp.

Figure 16 plots the optimal trade liberalization policy when reallocation effort is endogenous. Each panel considers one of our scenarios for households. For comparison, we plot the optimal tax with exogenous effort in each scenario. In this case, endogenous effort does not alter by much the trajectory of optimal policy.

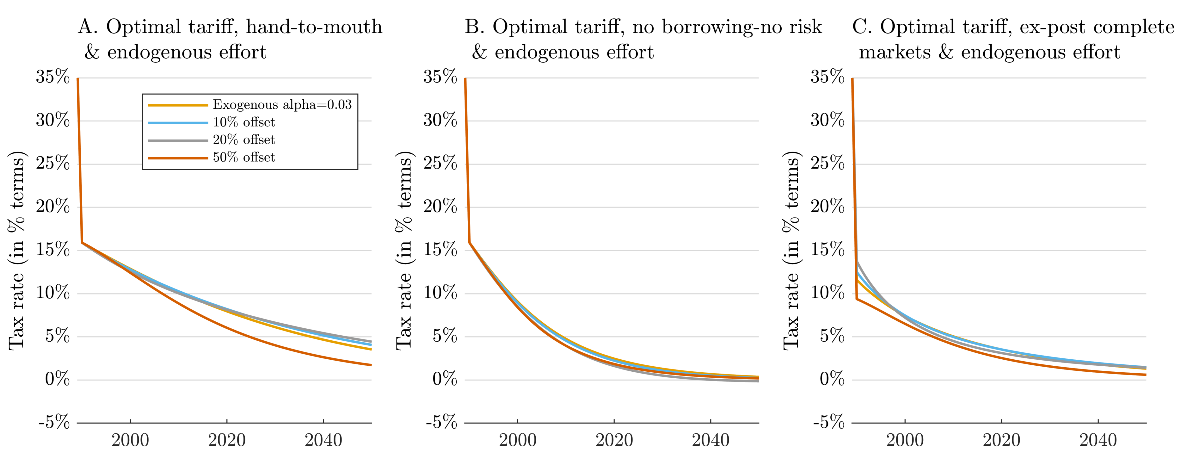

FIGURE 16: OPTIMAL TARIFF PATH FOR COLOMBIA'S TRADE LIBERALIZATION WHEN REALLOCATION EFFORT IS ENDOGENOUS. Panel A considers hand-to-mouth households. Panel B considers households that share reallocation risk but cannot borrow or save outside their islands. Panel C considers the case of ex-post complete markets.

6 Concluding remarks

This paper explored how gradualism mediates the gains from trade, technological change, and reforms. We argued that gradual changes have less adverse distributional effects in the short run and justify the use of temporary taxes or gradual reforms. We provided formulas for the optimal path for taxes in response to technological change, trade, or during a reform.

We applied our theory to studying the decline in routine jobs, the rise in Chinese import competition in the US, and Colombia's trade liberalization. A version of our model calibrated to match the short run income declines experienced by some workers as a result of the automation of routine jobs or rising import competition from China suggest that there are sizable gains from gradualism that justify temporary taxes of the order of 10%. Our formulas also suggest that the swift trade liberalization of 1990 in Colombia can only be justified in an scenario where workers can reallocate at a rate of 20% per year—an order of magnitude of what we estimate for the US.

The fact that short run taxes on automation and trade are desirable does not mean that the US economy did not benefit from rapid advances in Chinese exporting capabilities or the development of automation technologies. In both scenarios, we show that the gains from technological gradualism are negative, even in the absence of government policy. From a welfare point of view, rapid advances in China and in the development of automation technologies were a welcomed force. It is just that we could have made things better by taxing these technologies in the short run, easing the transition for displaced workers, and use the extra tax revenue.

Our formulas show that the desirability of taxes and the gains from technological gradualism depend on the extent to which disrupted households cut their consumption during a period of adjustment. Most of the existing literature focuses on estimating the impact of trade and technological disruptions on income. From a policy perspective, understanding how these disruptions affect consumption seems even more important, and a natural question for future research.

One interesting extension of our theory involves a case with congestion in reallocation; for example, because retraining a vast number of people in a single period might be subject to aggregate diminishing returns. Though we believe this offers an important rationale for gradualism, we did not explore the implications of congestion in our empirical applications.

References

Acemoglu, Daron and David Autor (2011): *Skills, Tasks and Technologies: Implications for Employment and Earnings*, Elsevier, vol. 4 of *Handbook of Labor Economics*, chap. 12, 1043–1171.

Acemoglu, Daron, Andrea Manera, and Pascual Restrepo (2020): "Does the US Tax Code Favor Automation?" *Brookings Papers on Economic Activity*, 231–285.

Acemoglu, Daron and Pascual Restrepo (2020): "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 128, 2188–2244.

——— (forthcoming): "Tasks, Automation, and the Rise in US Wage Inequality," *Econometrica*.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2021): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 89, 45–86.

Adão, Rodrigo, Martin Beraja, and Nitya Pandalai-Nayar (2021): "Fast and Slow Technological Transitions," Tech. rep., MIT, Mimeo.

Alvarez, Fernando and Robert Shimer (2011): "Search and Rets Unemployment," *Econometrica*, 79, 75–122.

Antràs, Pol, Alonso de Gortari, and Oleg Itskhoki (2017): "Globalization, inequality and welfare," *Journal of International Economics*, 108, 387–412.

Autor, David and David Dorn (2013): "The Growth of Low-skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97.

Autor, David, David Dorn, and Gordon H Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68.

Autor, David H, David Dorn, Gordon H Hanson, and Jae Song (2014): "Trade adjustment: Worker-level evidence," *The Quarterly Journal of Economics*, 129, 1799–1860.

Bai, Liang and Sebastian Stumpner (2019): "Estimating US Consumer Gains from Chinese Imports," *American Economic Review: Insights*, 1, 209–24.

Beraja, Martin and Nathan Zorzi (2022): "Inefficient Automation," Working Paper 30154, National Bureau of Economic Research.

Bond, Eric W. and Jee-Hyeong Park (2002): "Gradualism in Trade Agreements with Asymmetric Countries," *The Review of Economic Studies*, 69, 379–406.

Broda, Christian and David E. Weinstein (2006): "Globalization and the Gains From Variety," *The Quarterly Journal of Economics*, 121, 541–585.

Brynjolfsson, Erik and Andrew McAfee (2014): *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, W.W. Norton & Company.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835.

Chaney, Thomas (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–21.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber (2011): "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101, 471–75.

Chetverikov, Denis, Bradley Larsen, and Christopher Palmer (2016): "IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade," *Econometrica*, 84, 809–833.

Chisik, Richard (2003): "Gradualism in free trade agreements: a theoretical justification," *Journal of International Economics*, 59, 367–397.

Cortes, Guido Matias (2016): "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data," *Journal of Labor Economics*, 34, 63–105.

Costinot, Arnaud and Andrés Rodríguez-Clare (2014): "Chapter 4 - Trade Theory with Numbers: Quantifying the Consequences of Globalization," in *Handbook of International Economics*, ed. by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, Elsevier, vol. 4 of *Handbook of International Economics*, 197–261.

Costinot, Arnaud and Iván Werning (2018): "Robots, Trade, and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation," Working Paper 25103, National Bureau of Economic Research.

——— (forthcoming): "Robots, Trade and Luddism: A Sufficient Statistic Approach to Optimal Technology Regulation," *The Review of Economic Studies*.

Diamond, Peter A and James A Mirrlees (1971): "Optimal taxation and public production I: Production efficiency," *The American economic review*, 61, 8–27.

Donald, Eric (2022): "Optimal Taxation with Automation: Navigating Capital and Labor's Complicated Relationship," Mimeo, Boston University.

Edwards, Sebastian (1994): "Reformas comerciales en América Latina," *Coyuntura Economica, Fedesarrollo*.

Edwards, Sebastián and Roberto Steiner (2008): *La revolución incompleta: Las reformas de Gaviria*, Editorial Norma.

Edwards, Sebastian and Sweder van Wijnbergen (1989): "Disequilibrium and structural adjustment," Elsevier, vol. 2 of *Handbook of Development Economics*, 1481–1533.

Eslava, Marcela, John Haltiwanger, Adriana Kugler, and Maurice Kugler (2013): "Trade and market selection: Evidence from manufacturing plants in Colombia," *Review of Economic Dynamics*, 16, 135–158, special issue: Misallocation and Productivity.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi (2022): "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade," *The Review of Economic Studies*, rdac020.

Goldberg, Pinelopi and Nina Pavcnik (2005): "Trade, wages, and the political economy of trade protection: evidence from the Colombian trade reforms," *Journal of International Economics*, 66, 75–105.

Goos, Maarten, Alan Manning, and Anna Salomons (2014): "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 104, 2509–26.

Grossman, Gene M and Elhanan Helpman (1994): "Protection for Sale," *The American Economic Review*, 84, 833–850.

Guerreiro, Joao, Sergio Rebelo, and Pedro Teles (2021): "Should Robots Be Taxed?" *The Review of Economic Studies*, 89, 279–311.

Head, Keith and Thierry Mayer (2014): "Gravity Equations: Workhorse,Toolkit, and Cookbook," Elsevier, vol. 4, chap. Chapter 3, 131–195.

Helpman, Elhanan (1997): *Politics and Trade Policy*, New York: Cambridge University Press.

Holmes, Thomas J. and John J. Stevens (2014): "An Alternative Theory of the Plant Size Distribution, with Geography and Intra- and International Trade," *Journal of Political Economy*, 122, 369–421.

Kaplan, Greg, Giovanni Violante, and Justin Weidner (2014): "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity.*

Karp, Larry and Thierry Paul (1994): "Phasing in and Phasing Out Protectionism with Costly Adjustment of Labour," *The Economic Journal*, 104, 1379–1392.

Lucas, Robert E and Edward C Prescott (1974): "Equilibrium search and unemployment," *Journal of Economic Theory*, 7, 188–209.

Mussa, Michael (1984): "The Adjustment Process and the Timing of Trade Liberalization," Working Paper 1458, National Bureau of Economic Research.

Naito, Hisahiro (1999): "Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency," *Journal of Public Economics*, 71, 165–188.

Pierce, Justin R. and Peter K. Schott (2016): "The Surprisingly Swift Decline of US Manufacturing Employment," *American Economic Review*, 106, 1632–62.

Rodrik, Rani (1995): "Trade and industrial policy reform," Elsevier, vol. 3 of *Handbook of Development Economics*, 2925–2982.

Saez, Emmanuel (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 68, 205–229.

Saez, Emmanuel and Stefanie Stantcheva (2016): "Generalized social marginal welfare weights for optimal tax theory," *American Economic Review*, 106, 24–45.

Thuemmel, Uwe (2018): "Optimal Taxation of Robots," Working Paper 7317, CESifo.

Tsyvinski, Aleh and Nicolas Werquin (2017): "Generalized compensation principle," Tech. rep., National Bureau of Economic Research.

US Census (2022): "Quarterly E-Commerce Report," Data retrieved from Census Monthly Retail Trade Indicators, `https://www.census.gov/retail/index.html`.

Table 1: Calibration for the decline in routine jobs over 1985–2007.

Panel I. Islands and technology

Occupation	*Data & moments, Cortes (2016)*: Employment share 1985	Wage decline 85–07	Incidence	*Estimated objects*: Size disrupted islands, $\nu_{o(x)}$	Share disrupted, $s_{o(x)}$	S−curve parameters $\{\pi_x, \kappa_x\}$
Clerical jobs	10%	-14.4%	86.1%	5.6%	56.3%	$\{0.3, 1.0615\}$
Production jobs	18.5%	-30.1%	71.1%	16.0%	86.6%	$\{0.3, 1.7104\}$
Sales jobs	5%	-11.9%	18%	2.46%	49.1%	$\{0.3, 11.325\}$
Handling jobs	4%	-40.8%	96.4%	3.98%	99.4%	$\{0.3, 1.3335\}$

Panel II. Elasticities, reallocation rate, and households

Elasticity of substitution	$\sigma = 0.85$	From literature on polarization (see Goos et al., 2014)
Reallocation rate per year	$\alpha_0 = 2.5\%$	Calibrated to match average incidence of 71%
Inverse elasticity of intertemporal substitution	$\gamma = 2$	Standard macro calibration.
Discount rate and interest rate	$r = \rho = 5\%$	Standard macro calibration.
Initial assets	0	Low median liquid assets in US Survey of Consumer Finances

Notes: The table summarizes the data used to calibrate the model to match the wage decline in routine jobs and the resulting parameters. The Employment shares of routine occupations come from Acemoglu and Autor (2011); their wage decline from 1985–2007 from Cortes (2016); and the incidence of the wage decline also from Cortes (2016). The scale parameter of the S-curve h_x in equation (12) is not reported because it has no clear interpretation. Section 3 describes the calibration approach and data in detail.

Table 2: Calibration for the China Shock 1991–2007.

Panel I. Islands and technology

SIC code and industry	Data & moments, Autor et al. (2013), Autor et al. (2014)			Estimated objects		
	Value-added share 1991	Normalized import share 91–07 (pp)	Import Penetration 91–07 (pp)	Size disrupted islands, $\nu_{i(x)}$	Share disrupted, $s_{i(x)}$	S−curve parameters $\{\pi_x, \kappa_x\}$
20 Food & Kindred Products	1.77%	0.87	0.48	0.01%	0.74%	{0.5,3.9768}
21 Tobacco Products	0.29%	0.02	0.02	0.00%	0.01%	{0.5,14.217}
22 Textile Mill Products	0.32%	2.8	1.99	0.01%	2.38%	{0.5,4.3931}
23 Apparel	0.43%	35.97	21.76	0.13%	30.57%	{0.5,3.4398}
24 Lumber & Wood Products	0.36%	6.74	4.05	0.02%	5.72%	{0.5,2.3346}
25 Furniture & Fixtures	0.28%	39.69	27.88	0.09%	33.73%	{0.5,4.5851}
26 Paper & Allied Products	0.76%	2.75	1.83	0.02%	2.34%	{0.5,2.1375}
27 Printing & Publishing	1.31%	1.07	1.03	0.01%	0.91%	{0.5,1.6949}
28 Chemical & Allied Products	1.95%	1.94	1.58	0.03%	1.65%	{0.5,2.8831}
29 Petroleum & Coal Products	0.35%	0.54	0.13	0.00%	0.46%	{0.5,3.4844}
30 Rubber & Plastics Products	0.63%	10.53	7.95	0.06%	8.95%	{0.5,1.4856}
31 Leather & Leather Products	0.06%	108.38	58.44	0.05%	92.11%	{0.5,0.5088}
32 Stone, Clay, & Glass Products	0.43%	8.06	6.53	0.03%	6.85%	{0.5,1.0715}
33 Primary Metal Industries	0.66%	9.29	4.95	0.05%	7.90%	{0.5,3.0595}
34 Fabricated Metal Products	1.03%	8.69	6.37	0.07%	7.38%	{0.5,2.5377}
35 Industrial Machinery	1.67%	24.32	19.33	0.34%	20.67%	{0.5,2.9514}
36 Electronic Equipment	1.36%	36.04	25.96	0.41%	30.63%	{0.5,1.9978}
37 Transportation Equipment	1.88%	2.44	1.32	0.04%	2.07%	{0.5,2.5975}
38 Instruments & Related	1.04%	4.46	4.26	0.04%	3.79%	{0.5,0.99679}
39 Miscellaneous Manufacturing	0.26%	70.49	43.05	0.15%	59.91%	{0.5,0.9271}

Panel II. Elasticities, reallocation rate, and households

Elasticity of substitution	$\sigma = 2$	From Broda and Weinstein (2006)
Reallocation rate per year	$\alpha_0 = 1.8\%$	Calibrated to match incidence regressions in Autor et al. (2014)
Inverse elasticity of intertemporal substitution	$\gamma = 2$	Standard macro calibration.
Discount rate and interest rate	$r = \rho = 5\%$	Standard macro calibration.
Initial assets	0	Low median liquid assets in US Survey of Consumer Finances

Notes: The table summarizes the data used to calibrate the model to match the China Shock and the resulting parameters. Industry value added shares come from the NBER-CES, and are adjusted using aggregate data from the BEA-BLS integrated industry accounts to recognize the fact that the NBER-CES does not remove intermediate services from value added. Normalized import shares and import penetration measures come from Autor et al. (2013) and Autor et al. (2014). The scale parameter of the S-curve h_x in equation (12) is not reported because it has no clear interpretation. Section 4 describes the calibration approach and data in detail.

Table 3: Calibration for Colombia's trade liberalization.

Panel I. Islands and technology

SIC code and industry	*Data & moments, Eslava et al. (2013), DNP* Value-added share 1989	Effective tariff 1989	Percent decline in effective tariff	Change normalized import shares 89–02	*Estimated objects* Size disrupted islands, $\nu_{i(x)}$	Share disrupted, $s_{i(x)}$
311 Food products	1.59%	48.86%	158.71%	27.6 pp	0.40%	22.19%
312 Food	1.59%	30.72%	91.34%	9.1 pp	0.18%	10.07%
313 Beverages	2.42%	41.53%	95.03%	0.4 pp	0.01%	0.37%
314 Tobacco	0.43%	32.10%	80.42%	11.9 pp	0.06%	12.92%
321 Textiles	2.01%	44.61%	111.89%	28.9 pp	0.58%	24.92%
322 Apparel	0.58%	52.73%	116.54%	2.2 pp	0.01%	1.66%
323 Leather products	0.14%	32.13%	70.53%	14.9 pp	0.03%	16.11%
324 Shoes	0.23%	55.14%	126.10%	13.9 pp	0.03%	10.15%
331 Wood products	0.15%	38.15%	84.11%	12.9 pp	0.02%	12.44%
332 Furniture	0.10%	19.30%	64.67%	16.0 pp	0.03%	22.93%
341 Paper products	0.72%	28.10%	61.72%	10.2 pp	0.10%	12.03%
342 Printing and publishing	0.59%	28.19%	73.49%	6.4 pp	0.05%	7.53%
352 Chemical products	1.40%	25.37%	48.22%	48.6 pp	0.98%	60.58%
354 Oil and coal derivatives	0.11%	21.97%	43.35%	41.8 pp	0.07%	56.32%
355 Rubber	0.31%	29.41%	65.44%	58.0 pp	0.24%	66.29%
356 Plastic products	0.55%	35.43%	93.49%	41.9 pp	0.27%	42.49%
361 Clay products	0.15%	41.27%	93.38%	6.8 pp	0.01%	6.16%
362 Glass	0.25%	21.86%	45.24%	17.3 pp	0.07%	23.34%
369 Mineral products	0.89%	19.32%	48.11%	1.8 pp	0.03%	2.63%
381 Metal products	0.65%	23.63%	59.76%	36.7 pp	0.35%	47.61%
382 Machinery (exc. electric)	0.34%	15.25%	32.95%	14.7 pp	0.09%	23.48%
383 Electronics	0.73%	21.56%	45.87%	107.6 pp	0.84%	100.00%
384 Transportation equipment	1.09%	40.92%	101.65%	73.7 pp	0.84%	67.67%
385 Instruments	0.17%	22.11%	36.90%	98.2 pp	0.20%	100.00%
390 Miscellaneous products	0.22%	24.84%	63.68%	62.4 pp	0.20%	78.62%

Panel II. Elasticities, reallocation rate, and households

Elasticity of substitution	$\sigma = 2$	Imputed from China-Shock application and Broda and Weinstein (2006)
Reallocation rate per year	$\alpha_0 = 3\%$	Matches decline in industry premium in Goldberg and Pavcnik (2005)
Inverse elasticity of intertemporal substitution	$\gamma = 2$	Standard macro calibration.
Discount rate and interest rate	$r = \rho = 5\%$	Standard macro calibration.
Initial assets	0	Imputed from China-Shock application

Notes: The table summarizes the data used to calibrate the model to match Colombia's trade liberalization and the resulting parameters. Industry value added shares come from the *Departamento Nacional De Planeacion*, DNP. Effective tariffs come from the *Ministerio de Comercio*, and are described in detail in Eslava et al. (2013). The change in import shares before and after the reform come from the *Departamento Nacional De Planeacion*, DNP. We compute the 1989 level of imports as an average over 1985–1989 and the post reform level as an average over 1998–2002. We exclude industries 351 "industrial chemicals", 353 "oil refined products", 371 "steel and iron" and 372 "primary metals" from the analysis because they was already liberalized prior to 1990 and experienced no increase in import penetration since then. Section 5 describes the calibration approach and data in detail.

Online Appendix to "Optimal Gradualism"

Nils Lehr and Pascual Restrepo

October 23, 2022

A.1 Proofs for Section 1

Proof of Proposition 1:. Suppose that $k_{x,t} > 0$. We verify this is the case at the end of the proof. Firms in island x must be indifferent between producing with workers or producing using the new technology. This implies $w_{x,t} = (1 + \tau_{x,t})/A_{x,t}$.

We now show that wages across undisrupted islands are equalized. Assumption 1 ensures that this holds at time 0. Each period, a flow $\alpha \cdot (1 - \ell_t)$ of workers joins these islands. Directed search implies that this flow of workers is allocated in a way that preserves the equality of wages. In particular, this is the case so long as new workers are allocated in proportion to island populations at the time of their arrival. This shows that, under Assumption 1, directed search is not strictly necessary. Random search as in Alvarez and Shimer (2011) suffices.

The expression pinning down the common wage w_t comes from the fact that we have normalized the price of the final good to 1, which implies that island wages lie along the iso-cost curve $1 = c^f\left(\{w_{x,t}\}_{x \in \mathcal{D}}, w_t\right)$. Note that this equation implies that w_t is implicitly a function of the vector of after tax productivities $\{(1 + \tau_{x,t})/A_{x,t}\}$.

To derive the expression for output and new technology utilization we use Shepard's lemma, which implies

$$y_{x,t} = y_t \cdot c^f_x.$$

Adding across undisrupted islands, yields

$$\ell_t = y_t \cdot \sum_{x \notin \mathcal{D}} c^f_x = y_t \cdot c^f_w\left(\{w_{x,t}\}_{x \in \mathcal{D}}, w_t\right),$$

which after rearrangement yields the expression for output. On the other hand, for disrupted islands we have

$$\ell_{x,t} + k_{x,t} = y_t \cdot c^f_x\left(\{w_{x,t}\}_{x \in \mathcal{D}}, w_t\right).$$

Substituting the expression for output and rearranging yields the expression for $k_{x,t}$ in the proposition.

To conclude, we show that $k_{x,t} > 0$ as claimed in the proposition (and initially in the proof). Note that this is needed to ensure that capital is actually used and that firms are therefore indifferent between hiring workers or capital. Assumption 2 implies

$$\frac{c_x^f\left(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t\right)}{c_w^f\left(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t\right)} > \frac{c_x^f\left(\{\bar{w}\}_{x\in\mathcal{D}}, \bar{w}\right)}{c_w^f\left(\{\bar{w}\}_{x\in\mathcal{D}}, \bar{w}\right)} = \frac{\ell_{x,0}}{\ell_0} \geq \frac{\ell_{x,t}}{\ell_t}.$$

Rearranging this inequality yields $k_{x,t} > 0$. ■

Proof of Proposition 2:. The first part follows from the definition of a complete market. The second and third parts follow from an envelope logic. Workers in undisrupted island can set $\alpha_x = 0$ and achieve the exact consumption path of workers in disrupted islands, since their wages are higher (or equal) at all points in time. They can then consume the excess savings generated with this policy, which implies that they can always achieve a strictly higher utility than disrupted households. Because they are making optimal decisions, we must have $U_0 > U_{x,0}$, as wanted. ■

A.2 Proofs for Section 2

This section provides proofs for Lemma 1 and Propositions 3, 4, and A3.

Proof of Lemma 1. We first derive the expression for the change in tax revenue in (1). We prove this in a slightly more general case. For simplicity, we ignore time subscripts. Suppose there are different types of labor indexed by ℓ_j, different (untaxed) intermediate goods y_i, and different types of (taxed) capital k_m. Tax revenue is given by

$$T = \sum_m \tau_m \cdot \frac{k_m}{A_m},$$

which implies

$$dT = \sum_m \tau_m \cdot \frac{dk_m}{A_m} + \sum_m \frac{k_m}{A_m} \cdot d\tau_m.$$

Each producer f operates some CRS production function. Thus, we have

$$0 = \max_{y^f, y_i^f, k_m^f, \ell_j^f} y^f + \sum_i y_i^f \cdot p_i - \sum_m \frac{k_m^f}{A_m} \cdot (1 + \tau_m) - \sum_j \ell_j^f \cdot w_j.$$

The envelope theorem then implies

$$0 = \sum_i y_i^f \cdot dp_i - \sum_m \frac{k_m^f}{A_m} \cdot d\tau_m - \sum_j \ell_j^f \cdot dw_j.$$

Adding this across all producers, we obtain

$$\sum_m \frac{k_m}{A_m} \cdot d\tau_m = -\sum_j \ell_j \cdot dw_j.$$

Plugging into the expression for dT yields a general version of the formula in the lemma:

$$dT = \sum_m \tau_m \cdot \frac{dk_m}{A_m} - \sum_j \ell_j \cdot dw_j.$$

We now turn to the change in welfare. The welfare function is given by

$$W_0 = \sum_{x \in \mathcal{D}} \ell_{x,0} \cdot \mathcal{W}\left(\max_\alpha \mathcal{U}_x(\{w_{x,t} + T_t, w_t + T_t\}_{t=0}^{\infty}; \alpha)\right) + \ell_0 \cdot \mathcal{W}\left(\mathcal{U}(\{w_t + T_t\}_{t=0}^{\infty})\right).$$

Recall that $\mathcal{U}_x(\{w_{x,t}+T_t, w_t+T_t\}_{t=0}^{\infty}; \alpha)$ gives the maximum utility that a displaced household can achieve. The envelope theorem implies that, so long as $c_{x,t} > 0$, an infinitesimal change in $w_{x,t} + T_t$ changes $\mathcal{U}_x(\{w_{x,t} + T_t, w_t + T_t\}_{t=0}^{\infty}; \alpha)$ by

$$d\mathcal{U}_x(\{w_{x,t} + T_t, w_t + T_t\}_{t=0}^{\infty}; \alpha) = P_{x,t} \cdot e^{-\rho t} \cdot u'(c_{x,t}) \cdot (dw_{x,t} + dT_t).$$

This is because optimizing households weakly prefer consuming the additional income to any other use. After all, households could always reduce their consumption but they choose not to do that. The right hand side gives the expected marginal utility of consuming this extra income.

Likewise, so long as $c_{x,t_r,t} > 0$, an infinitesimal change in $w_t + T_t$ changes $\mathcal{U}_x(\{w_{x,t} + T_t, w_t + T_t\}_{t=0}^{\infty}; \alpha)$ by

$$\begin{aligned} d\mathcal{U}_x(\{w_{x,t} + T_t, w_t + T_t\}_{t=0}^{\infty}; \alpha) = & \int_0^t e^{-\rho t} \cdot u'(c_{x,t_r,t}) \cdot (dw_t + dT_t) \cdot \cdot \alpha_x \cdot e^{-\alpha_x t_r} \cdot dt_r \\ = & (1 - P_{x,t}) \cdot e^{-\rho t} \cdot \mathbb{E}[u'(c_{x,t_r,t}) | t_r \leq t] \cdot (dw_t + dT_t), \end{aligned}$$

where this expression integrates over all potential histories in which the household is at the undisrupted island at time t and can consume this extra income.

Finally, so long as $c_t > 0$, an infinitesimal change in $w_t + T_t$ changes $\mathcal{U}(\{w_t + T_t\}_{t=0}^{\infty})$ by

$$d\mathcal{U}(\{w_t + T_t\}_{t=0}^{\infty}) = e^{-\rho t} \cdot u'(c_t) \cdot (dw_t + dT_t).$$

Combining these results yields the formula in equation (2).

Note that the lemma also applies when there is endogenous reallocation effort. This is because changes in reallocation effort have a second order effect on $U_{x,0}$ (households are optimizing with respect to α_x). ■

Proof of Proposition 3. The proof follows Costinot and Werning (forthcoming). Using the definition of the χ's, we can write the change in welfare following a variation as

$$dW_0 = \int_0^{\infty} \left[\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \chi_{x,t} \cdot (dw_{x,t} + dT_t) + \ell_t \cdot \chi_t \cdot (dw_t + dT_t) \right] dt. \tag{A1}$$

Now, consider a change in taxes that only changes $k_{x',t}$ by $dk_{x',t}$ but keeps the utilization of all other types of capital unchanged. At a social optimum, this variation cannot affect welfare. Thus:

$$\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \chi_{x,t} \cdot dw_{x,t} + \ell_t \cdot \chi_t \cdot dw_t + \bar{\chi}_t \cdot dT_t = 0.$$

Using the fact that $dT_t = -\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot dw_{x,t} - \ell_t \cdot dw_t + \tau_{x',t} \cdot \frac{dk_{x',t}}{A_{x',t}}$, we obtain

$$\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot (\chi_{x,t} - \bar{\chi}_t) \cdot dw_{x,t} + \ell_t \cdot (\chi_t - \bar{\chi}_t) \cdot dw_t + \bar{\chi}_t \cdot \tau_{x',t} \cdot \frac{dk_{x',t}}{A_{x',t}} = 0.$$

We can use this expression to solve for $\tau_{x',t}$ as

$$\tau_{x',t} = \sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \left(\frac{\chi_{x,t}}{\bar{\chi}_t} - 1 \right) \cdot \left(-\frac{dw_{x,t}}{dk_{x',t}} \cdot A_{x',t} \right) + \ell_t \cdot \left(\frac{\chi_t}{\bar{\chi}_t} - 1 \right) \cdot \left(-\frac{dw_t}{dk_{x',t}} \cdot A_{x',t} \right).$$

Using the fact that $m_{x,t} = k_{x,t}/A_{x,t}$, this can be rewritten as equation (3) ■

Proof of Proposition 4. As before, we will use Lemma 1. Suppose we are at an optimum. The optimum is described by some sequence of capital utilized by island. Consider a reform that changes $k_{x',t}$ by $dk_{x',t}$ but leaves all other $k_{x,s}$ unchanged. This reform will necessarily change α_x by $d\alpha_x$ and, because the reform kept $k_{x,s}$ fixed for all other x, s, it will necessarily change wages at all points in time and islands.

Define the direct effect of the reform as the effect on welfare through wages and tax revenue holding all α_x constant. This then triggers an indirect effect via changes in α_x which affect wages and revenue at all other time periods.

For any variable $a_{x,s}$ we denote by $d_k a_{x,s}$ the direct effect of the reform—i.e., the change induced by $k_{x',t}$—and by $d_\alpha a_{x,s}$ the indirect effect—i.e., the change induced by changes in α_x.

Let's first consider the indirect effects and the determination of $d\alpha_x$. The first-order condition for α_x is

$$\kappa'(\alpha_x) = \mathcal{U}_{x,\alpha}.$$

Totally differentiating this equation we get

$$\kappa''(\alpha_x) \cdot d\alpha_x = \sum_{x''} \theta_{x,x''} \cdot d\alpha_{x''} + \mathcal{U}_{x,\alpha,d,t} \cdot (d_k w_{x,t} + d_k T_t) \cdot dt + \mathcal{U}_{x,\alpha,r,t} \cdot (d_k w_t + d_k T_t) \cdot dt. \tag{A2}$$

In this equation, $\theta_{x,x''}$ gives the effect of changing $d\alpha_{x''}$ on $\mathcal{U}_{x,\alpha}$ via changes in wages and tax revenue over time (this object also has to be computed holding $k_{x,s}$ constant for all x, s).

The equation also shows that the direct effects of the reform also alter $\mathcal{U}_{x,\alpha}$, but these effects are "of the order of" dt, since the direct effect only changes wages and tax revenue at a point in time t.

Equation (A2) is a system of equations that can be solved as

$$d\alpha_x = \sum_{x''} \varepsilon_{x,x''} \cdot (\mathcal{U}_{x'',\alpha,d,t} \cdot (d_k w_{x'',t} + d_k T_t) + \mathcal{U}_{x'',\alpha,r,t} \cdot (d_k w_t + d_k T_t)) \cdot dt \tag{A3}$$

The $\varepsilon_{x,x''}$ tell us how changes in the incentives to reallocate in island x'' affect reallocation rates from island x, holding $k_{x,s}$ constant.

Let $\mu_x \cdot \ell_{x,0}$ denote the welfare gains from increasing the reallocation rate out of island x. This object must also be computed holding $k_{x,s}$ constant. It captures all the indirect changes in wages and tax revenue generated by changes in reallocation rates.

We can compute the welfare gains from the reform that changes $k_{x',t}$ by $dk_{x,t}$ and leaves all other $k_{x,s}$ unchanged as

$$\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \chi_{x,t} \cdot (d_k w_{x,t} + d_k T_t) \cdot dt + \ell_t \cdot \chi_t \cdot (d_k w_t + d_k T_t) \cdot dt + \sum_{x \in \mathcal{D}} \mu_x \cdot \ell_{x,0} \cdot d\alpha_x = 0.$$

Note that there is a dt multiplying the welfare effects of the direct effect of the reform via wages and tax revenue, since this only accrue in an instant of time. This implies that both the direct and indirect effects are "of the order of" dt.

Using the formula for $d\alpha_x$ in equation (A3) we get

$$\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi_{x,t}\cdot(d_k w_{x,t}+d_k T_t)\cdot dt+\ell_t\cdot\chi_t\cdot(d_k w_t+d_k T_t)\cdot dt$$
$$+\sum_{x\in\mathcal{D}}\mu_x\cdot\ell_{x,0}\cdot\sum_{x''}\varepsilon_{x,x''}\cdot(\mathcal{U}_{x'',\alpha,d,t}\cdot(d_k w_{x'',t}+d_k T_t)+\mathcal{U}_{x'',\alpha,r,t}\cdot(d_k w_t+d_k T_t))\cdot dt=0.$$

At this point, we can cancel the dt. It makes intuitive sense that this should cancel. The direct effect of a reform at time t on welfare should be of the order of dt. The direct effect of a reform at time t on reallocation rates should also be of the order of dt.

We can rearrange the above equation as

$$\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi_{x,t}\cdot(d_k w_{x,t}+d_k T_t)+\ell_t\cdot\chi_t\cdot(d_k w_t+d_k T_t)$$
$$+\sum_{x}\left(\sum_{x''\in\mathcal{D}}\mu_{x''}\cdot\ell_{x'',0}\cdot\varepsilon_{x'',x}\right)(\mathcal{U}_{x,\alpha,d,t}\cdot(d_k w_{x,t}+d_k T_t)+\mathcal{U}_{x,\alpha,r,t}\cdot(d_k w_t+d_k T_t))=0.$$

We can now define the social value of increasing income at different islands as

$$\chi^{\text{end}}_{x,t}=\chi_{x,t}+\sum_{x''\in\mathcal{D}}\frac{\ell_{x'',0}}{\ell_{x,t}}\cdot\mu_{x''}\cdot\varepsilon_{x'',x}\cdot\mathcal{U}_{x,\alpha,d,t},$$
$$\chi^{\text{end}}_{t}=\chi_{t}+\sum_{x,x''\in\mathcal{D}}\frac{\ell_{x'',0}}{\ell_{t}}\cdot\mu_{x''}\cdot\varepsilon_{x'',x}\cdot\mathcal{U}_{x,\alpha,r,t},$$
$$\bar{\chi}_t=\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi^{\text{end}}_{x,t}+\ell_t\cdot\chi^{\text{end}}_t.$$

Following the same steps as in the proof of Proposition 3, we get

$$\tau_{x',t}=\sum_{x\in\mathcal{D}}\frac{\ell_{x,t}\cdot w_{x,t}}{m_{x',t}}\cdot\left(\frac{\chi^{\text{end}}_{x,t}}{\bar{\chi}_t}-1\right)\cdot\left(-\frac{d_k\ln w_{x,t}}{d_k\ln k_{x',t}}\right)+\frac{\ell_t\cdot w_t}{m_{x',t}}\cdot\left(\frac{\chi^{\text{end}}_{t}}{\bar{\chi}_t}-1\right)\cdot\left(-\frac{d_k\ln w_{t}}{d_k\ln k_{x',t}}\right)$$

This derivation also clarifies that $\frac{d_k\ln w_{x,t}}{d_k\ln k_{x',t}}$ and $\frac{d_k\ln w_{t}}{d_k\ln k_{x',t}}$ are partial derivatives: they correspond to the change in wages given a change in capital holding reallocation constant, which is not what one would estimate from the data.

Computing μ_x: We now compute the change in welfare driven by a change in $d\alpha_{x'}$ holding $k_{x,s}$ constant at all periods. Equation (A1) implies that these welfare gains are given by

$$\mu_{x'} \cdot \ell_{x',0} \cdot d\alpha_{x'} = \int_0^\infty \left[\sum_{x\in\mathcal{D}} \ell_{x,s} \cdot \chi_{x,s} \cdot d_\alpha w_{x,s} + \ell_s \cdot \chi_s \cdot d_\alpha w_s + \bar{\chi}_s \cdot d_\alpha T_s) \right] ds.$$

Moreover, we have

$$\begin{aligned}
d_\alpha w_{x,s} =& \frac{\partial w_{x,s}}{\partial \ell_{x',s}} \cdot (-s \cdot e^{-\alpha_{x'} s}) \cdot \ell_{x',0} \cdot d\alpha_{x'} \\
d_\alpha w_s =& \frac{\partial w_s}{\partial \ell_{x',s}} \cdot (-s \cdot e^{-\alpha_{x'} s}) \cdot \ell_{x',0} \cdot d\alpha_{x'} \\
d_\alpha T_s =& - \sum_{x\in\mathcal{D}} \ell_{x,s} \cdot \frac{\partial w_{x,s}}{\partial \ell_{x',s}} \cdot (-s \cdot e^{-\alpha_{x'} s}) \cdot \ell_{x',0} \cdot d\alpha_{x'} - \ell_s \cdot \frac{\partial w_s}{\partial \ell_{x',s}} \cdot (-s \cdot e^{-\alpha_{x'} s}) \cdot \ell_{x',0} \cdot d\alpha_{x'}.
\end{aligned}$$

Note that these are partial derivatives since we are interested on the effect of $\alpha_{x'}$ on wages and tax revenues holding all other α_x and $k_{x,s}$ constant. The last line uses Lemma 1 and the fact that the variation we are considering keeps all quantities constant (except for α_x). Plugging in the expression for $\mu_{x'}$ yields the formula in equation (5). ■

A.2.1 Deriving formulas for $\mathcal{U}_{x,\alpha}$, $\mathcal{U}_{x,\alpha,d,t}$, $\mathcal{U}_{x,\alpha,r,t}$.

Hand-to-mouth: In this case, we have

$$\mathcal{U}_x = \int_0^\infty e^{-\rho t} \cdot [P_{x,t} \cdot u(w_{x,t} + T_t) + (1 - P_{x,t}) \cdot u(w_t + T_t)] \cdot dt.$$

Differentiating this with respect to α, and then with respect to wages at time t, we obtain:

$$\begin{aligned}
\mathcal{U}_{x,\alpha} =& \int_0^\infty e^{-\rho t} \cdot (t \cdot P_{x,t}) \cdot [u(w_t + T_t) - u(w_{x,t} + T_t)] \cdot dt, \\
\mathcal{U}_{x,\alpha,d,t} =& - (t \cdot P_{x,t}) \cdot \lambda_{x,d,t}, \\
\mathcal{U}_{x,\alpha,r,t} =& (t \cdot P_{x,t}) \cdot \lambda_{x,r,t}.
\end{aligned}$$

No borrowing and no transition risk: Let

$$c_{x,t} = P_{x,t} \cdot w_{x,t} + (1 - P_{x,t}) \cdot w_t + T_t$$

In this case, we have

$$\mathcal{U}_x = \int_0^\infty e^{-\rho t} \cdot u\left(c_{x,t}\right) \cdot dt.$$

Differentiating this with respect to α, and then with respect to wages at time t, we obtain:

$$\begin{aligned}
\mathcal{U}_{x,\alpha} &= \int_0^\infty e^{-\rho t} \cdot \left(t \cdot P_{x,t}\right) \cdot \left(w_t - w_{x,t}\right) \cdot u'\left(c_{x,t}\right) dt, \\
\mathcal{U}_{x,\alpha,d,t} &= -\left(t \cdot P_{x,t}\right) \cdot \lambda_{x,d,t} + e^{-\rho t} \cdot \left(t \cdot P_{x,t}\right) \cdot \left(w_t - w_{x,t}\right) \cdot u''\left(c_{x,t}\right) \cdot P_{x,t}, \\
\mathcal{U}_{x,\alpha,r,t} &= \left(t \cdot P_{x,t}\right) \cdot \lambda_{x,r,t} + e^{-\rho t} \cdot \left(t \cdot P_{x,t}\right) \cdot \left(w_t - w_{x,t}\right) \cdot u''\left(c_{x,t}\right) \cdot \left(1 - P_{x,t}\right).
\end{aligned}$$

Borrowing with transition risk: In this case there are no simple analytical expressions for $\mathcal{U}_{x,\alpha}$, $\mathcal{U}_{x,\alpha,d,t}$, $\mathcal{U}_{x,\alpha,r,t}$, nor a simple way of computing these objects numerically. For this reason, we do not analyze this scenario with endogenous reallocation effort.

Ex-post complete markets: Assume that $u(c) = c^{1-\gamma}/(1-\gamma)$ and let

$$h_{x,0} = a_{x,0} + \int_0^\infty e^{-rt} \cdot \left[P_{x,t} \cdot w_{x,t} + \left(1 - P_{x,t}\right) \cdot w_t + T_t\right] \cdot dt$$

denote the effective wealth of households in disrupted islands at time 0. We can solve analytically for $\mathcal{U}_x$ as

$$\mathcal{U}_x = \left[r - \frac{1}{\gamma}(r - \rho)\right]^{-\gamma} \cdot h_{x,0}^{1-\gamma}/(1-\gamma).$$

Differentiating this with respect to α, and then with respect to wages at time t, we obtain:

$$\begin{aligned}
\mathcal{U}_{x,\alpha} &= \left[r - \frac{1}{\gamma}(r - \rho)\right]^{-\gamma} \cdot h_{x,0}^{-\gamma} \cdot \int_0^\infty e^{-rt} \cdot \left(t \cdot P_{x,t}\right) \cdot \left(w_t - w_{x,t}\right) \cdot dt, \\
\mathcal{U}_{x,\alpha,d,t} &= -\left[r - \frac{1}{\gamma}(r - \rho)\right]^{-\gamma} \cdot h_{x,0}^{-\gamma} \cdot e^{-rt} \cdot \left(t \cdot P_{x,t}\right) - \gamma \cdot \frac{P_{x,t}}{h_{x,0}} \cdot \mathcal{U}_{x,\alpha}, \\
\mathcal{U}_{x,\alpha,r,t} &= \left[r - \frac{1}{\gamma}(r - \rho)\right]^{-\gamma} \cdot h_{x,0}^{-\gamma} \cdot e^{-rt} \cdot \left(t \cdot P_{x,t}\right) - \gamma \cdot \frac{1 - P_{x,t}}{h_{x,0}} \cdot \mathcal{U}_{x,\alpha}.
\end{aligned}$$

A.2.2 Temporary Assistance Programs

We now consider optimal policy for a government that can tax new technologies and also set up temporary assistance programs. We let households choose their level of work effort or work hours $n_{x,t}$. We also assume that flow utility is given by $u(c_{x,t} - \psi(n_{x,t}))$ so that there are no income effects. We assume ψ is a power function and that ε_ℓ is the constant elasticity of effort to wages. Without loss of generality we normalize initial effort to 1 at all islands.

With these modifications, the choice of work effort is static and given by

$$n_{x,t} = \left(\frac{w_{x,t} \cdot (1 - \mathcal{R}_t)}{\bar{w}}\right)^{\varepsilon_\ell}$$

at disrupted islands, and by

$$n_t = \left(\frac{w_t \cdot (1 - \mathcal{R}_t)}{\bar{w}}\right)^{\varepsilon_\ell}$$

at undisrupted islands. Moreover, the counterfactual work level of non participants in undisrupted islands is

$$n_t^* = \left(\frac{w_t}{\bar{w}}\right)^{\varepsilon_\ell}.$$

In what follows, define $\tilde{T}_t = T_t + \mathcal{R}_t \cdot w_t \cdot n_t^*$. Households' income at time t then becomes

$$(1 - \mathcal{R}_t) \cdot w_{x,t} \cdot n_{x,t} + \tilde{T}_t$$

for disrupted households, and

$$(1 - \mathcal{R}_t) \cdot w_t \cdot n_t + \tilde{T}_t$$

for households in undisrupted islands. This shows that disrupted households make more use of the assistance program, but because of imperfect targeting, households in disrupted islands will also make use of it.

Using this notation, we can write the condition for a balanced government budget as

$$\tilde{T}_t = \sum_{x \in \mathcal{D}} \tau_{x,t} \cdot \frac{k_{x,t}}{A_{x,t}} + \sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \mathcal{R}_t \cdot w_{x,t} \cdot n_{x,t} + \ell_t \cdot \mathcal{R}_t \cdot w_t \cdot n_t.$$

This shows that, in this formulation, temporary assistance programs are equivalent to having a time varying income tax $\mathcal{R}_t$.

Lemma A1 (Variations Lemma) *Consider a variation in taxes that induces a change in wages $dw_t, dw_{x,t}$, technology $dk_{x,t}$, tax revenue $d\tilde{T}_t$, and reallocation effort $d\alpha_x$ but keep*

$\mathcal{R}_t$ constant. This variation changes tax revenue by

(A4)
$$d\tilde{T}_t = -(1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\cdot\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot dw_{x,t} - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\cdot\ell_t\cdot n_t\cdot dw_t + \sum_x \tau_x\cdot\frac{dk_{x,t}}{A_{x,t}}$$

and social welfare by

$$dW_0 = \int_0^\infty \Bigg[\sum_{x\in\mathcal{D}}\ell_{x,0}\cdot g_x\cdot\Big(P_{x,t}\cdot\lambda_{x,d,t}\cdot((1-\mathcal{R}_t)\cdot n_{x,t}\cdot dw_{x,t}+d\tilde{T}_t) + (1-P_{x,t})\cdot\lambda_{x,r,t}\cdot((1-\mathcal{R}_t)\cdot n_t\cdot dw_t + d\tilde{T}_t)\Big) + \ell_0\cdot g\cdot\lambda_t\cdot((1-\mathcal{R}_t)\cdot n_t\cdot dw_t + d\tilde{T}_t)\Bigg]\cdot dt, \tag{A5}$$

where $\lambda_{x,d,t} = e^{-\rho t}\cdot u'(c_{x,t}-\psi(n_{x,t}))$ denotes the marginal utility of consumption at time t for households that have not reallocated, $\lambda_{x,r,t} = \mathbb{E}[e^{-\rho t}\cdot u'(c_{x,t_r,t}-\psi(n_{x,t}))|t_r\le t]$ denotes the average marginal utility of consumption among households that reallocated by time t, and $\lambda_t = e^{-\rho t}\cdot u'(c_t-\psi(n_t))$ denotes the marginal utility of consumption of non-disrupted households at time t.

Lemma A2 (Variations Lemma) *Consider a variation in the replacement rate $d\mathcal{R}_t$ that induces a change in wages $dw_t, dw_{x,t}$, tax revenue $d\tilde{T}_t$, and reallocation effort $d\alpha_x$ but keep $dk_{x,t}$ constant (by adjusting taxes on new technologies appropriately). This variation changes tax revenue by*

$$d\tilde{T}_t = -(1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\cdot\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot dw_{x,t} - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\cdot\ell_t\cdot n_t\cdot dw_t + \sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}\cdot\left(1-\varepsilon_\ell\frac{\mathcal{R}_t}{1-\mathcal{R}_t}\right)\cdot d\mathcal{R}_t + \ell_t\cdot n_t\cdot w_t\cdot\left(1-\varepsilon_\ell\frac{\mathcal{R}_t}{1-\mathcal{R}_t}\right)\cdot d\mathcal{R}_t \tag{A6}$$

and social welfare by

$$dW_0 = \int_0^\infty \Bigg[\sum_{x\in\mathcal{D}}\ell_{x,0}\cdot g_x\cdot\Big(P_{x,t}\cdot\lambda_{x,d,t}\cdot((1-\mathcal{R}_t)\cdot n_{x,t}\cdot dw_{x,t} - n_{x,t}\cdot w_{x,t}\cdot d\mathcal{R}_t + d\tilde{T}_t) + (1-P_{x,t})\cdot\lambda_{x,r,t}\cdot((1-\mathcal{R}_t)\cdot n_t\cdot dw_t - n_t\cdot w_t\cdot d\mathcal{R}_t + d\tilde{T}_t)\Big) + \ell_0\cdot g\cdot\lambda_t\cdot((1-\mathcal{R}_t)\cdot n_t\cdot dw_t - n_t\cdot w_t\cdot d\mathcal{R}_t + d\tilde{T}_t)\Bigg]\cdot dt, \tag{A7}$$

where $\lambda_{x,d,t} = e^{-\rho t} \cdot u'(c_{x,t} - \psi(n_{x,t}))$ denotes the marginal utility of consumption at time t for households that have not reallocated, $\lambda_{x,r,t} = \mathbb{E}[e^{-\rho t} \cdot u'(c_{x,t_r,t} - \psi(n_{x,t}))|t_r \leq t]$ denotes the average marginal utility of consumption among households that reallocated by time t, and $\lambda_t = e^{-\rho t} \cdot u'(c_t - \psi(n_t))$ denotes the marginal utility of consumption of non-disrupted households at time t.

Proof of Proposition 5. The proof follows Costinot and Werning (forthcoming). Using the definition of the χ's, we can write the change in welfare following a variation as

(A8)

$$dW_0 = \int_0^\infty \left[\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \chi_{x,t} \cdot ((1 - \mathcal{R}_t) \cdot n_{x,t} \cdot dw_{x,t} + d\tilde{T}_t) + \ell_t \cdot \chi_t \cdot ((1 - \mathcal{R}_t) \cdot n_t \cdot dw_t + d\tilde{T}_t) \right] dt.$$

Now, consider a change in taxes that only changes $k_{x',t}$ by $dk_{x',t}$ but keeps the utilization of all other types of capital unchanged. At a social optimum, this variation cannot affect welfare. Thus:

$$\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \chi_{x,t} \cdot (1 - \mathcal{R}_t) \cdot n_{x,t} \cdot dw_{x,t} + \ell_t \cdot \chi_t \cdot (1 - \mathcal{R}_t) \cdot n_t \cdot dw_t + \bar{\chi}_t \cdot d\tilde{T}_t = 0.$$

Using the expression for $d\tilde{T}_t$ in equation (A4), we obtain

$$\begin{aligned}
&\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot \chi_{x,t} \cdot (1 - \mathcal{R}_t) \cdot n_{x,t} \cdot dw_{x,t} \\
&+ \ell_t \cdot \chi_t \cdot (1 - \mathcal{R}_t) \cdot n_t \cdot dw_t \\
&+ \bar{\chi}_t \cdot \left(-(1 - \mathcal{R}_t \cdot (1 + \varepsilon_\ell)) \cdot \sum_{x \in \mathcal{D}} \ell_{x,t} \cdot n_{x,t} \cdot dw_{x,t} - (1 - \mathcal{R}_t \cdot (1 + \varepsilon_\ell)) \cdot \ell_t \cdot n_t \cdot dw_t + \tau_{x',t} \cdot \frac{dk_{x',t}}{A_{x',t}} \right) = 0.
\end{aligned}$$

We can use this expression to solve for $\tau_{x',t}$ as

$$\begin{aligned}
\tau_{x',t} = &\sum_{x \in \mathcal{D}} \ell_{x,t} \cdot n_{x,t} \cdot \left(\frac{\chi_{x,t}}{\bar{\chi}_t} \cdot (1 - \mathcal{R}_t) - (1 - \mathcal{R}_t \cdot (1 + \varepsilon_\ell)) \right) \cdot \left(-\frac{dw_{x,t}}{dk_{x',t}} \cdot A_{x',t} \right) \\
&+ \ell_t \cdot n_t \cdot \left(\frac{\chi_t}{\bar{\chi}_t} \cdot (1 - \mathcal{R}_t) - (1 - \mathcal{R}_t \cdot (1 + \varepsilon_\ell)) \right) \cdot \left(-\frac{dw_t}{dk_{x',t}} \cdot A_{x',t} \right).
\end{aligned}$$

Using the fact that $m_{x,t} = k_{x,t}/A_{x,t}$, this can be rewritten as

$$\text{(A9)} \qquad \tau_{x',t} = \sum_{x\in\mathcal{D}} \frac{\ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}}{m_{x',t}} \cdot \left(\frac{\chi_{x,t}}{\bar{\chi}_t}\cdot(1-\mathcal{R}_t) - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\right)\cdot\left(-\frac{\partial \ln w_{x,t}}{\partial \ln k_{x',t}}\right)$$
$$+ \frac{\ell_t\cdot n_t\cdot w_t}{m_{x',t}}\cdot\left(\frac{\chi_t}{\bar{\chi}_t}\cdot(1-\mathcal{R}_t) - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\right)\cdot\left(-\frac{\partial \ln w_t}{\partial \ln k_{x',t}}\right),$$

which can then be rearranged into the formula in (6).

We now turn to the formula for optimal replacement rates. Rewrite equation (A6) as

$$\text{(A10)} \qquad dW_0 = \int_0^\infty \Bigg[\sum_{x\in\mathcal{D}} \ell_{x,t}\cdot\chi_{x,t}\cdot\left((1-\mathcal{R}_t)\cdot n_{x,t}\cdot dw_{x,t} - n_{x,t}\cdot w_{x,t}\cdot d\mathcal{R}_t + d\tilde{T}_t\right)$$
$$+ \ell_t\cdot\chi_t\cdot\left((1-\mathcal{R}_t)\cdot n_t\cdot dw_t - n_t\cdot w_t\cdot d\mathcal{R}_t + d\tilde{T}_t\right)\Bigg]dt.$$

Consider a change in taxes that changes $d\mathcal{R}_t$ but keeps the utilization of capital unchanged. At a social optimum, this variation cannot affect welfare. Thus:

$$\sum_{x\in\mathcal{D}} \ell_{x,t}\cdot\chi_{x,t}\cdot\left((1-\mathcal{R}_t)\cdot n_{x,t}\cdot dw_{x,t} - n_{x,t}\cdot w_{x,t}\cdot d\mathcal{R}_t\right)$$
$$+ \ell_t\cdot\chi_t\cdot\left((1-\mathcal{R}_t)\cdot n_t\cdot dw_t - n_t\cdot w_t\cdot d\mathcal{R}_t\right) + \bar{\chi}_t\cdot d\tilde{T}_t = 0.$$

Using the expression for $d\tilde{T}_t$ in equation (A6) and rearranging terms yields

$$\text{(A11)} \qquad 0 = \sum_{x\in\mathcal{D}} \ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}\cdot\left(\frac{d\ln w_{x,t}}{d\mathcal{R}_t} - \frac{1}{1-\mathcal{R}_t}\right)\cdot\left(\frac{\chi_{x,t}}{\bar{\chi}_t}\cdot(1-\mathcal{R}_t) - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\right)$$
$$+ \ell_t\cdot n_t\cdot w_t\cdot\left(\frac{d\ln w_t}{d\mathcal{R}_t} - \frac{1}{1-\mathcal{R}_t}\right)\cdot\left(\frac{\chi_t}{\bar{\chi}_t}\cdot(1-\mathcal{R}_t) - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\right).$$

In the absence of GE effects from assistance programs, we have

$$0 = \sum_{x\in\mathcal{D}} \ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}\cdot\left(\frac{1}{1-\mathcal{R}_t}\right)\cdot\left(\frac{\chi_{x,t}}{\bar{\chi}_t}\cdot(1-\mathcal{R}_t) - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\right)$$
$$+ \ell_t\cdot n_t\cdot w_t\cdot\left(\frac{1}{1-\mathcal{R}_t}\right)\cdot\left(\frac{\chi_t}{\bar{\chi}_t}\cdot(1-\mathcal{R}_t) - (1-\mathcal{R}_t\cdot(1+\varepsilon_\ell))\right).$$

which can be rearranged as the usual optimal tax formula

$$\text{(A12)} \qquad \frac{\mathcal{R}_t}{1-\mathcal{R}_t} = \frac{1}{\varepsilon_\ell}\cdot\frac{\sum_{x\in\mathcal{D}} \ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}\cdot\left(1-\frac{\chi_{x,t}}{\bar{\chi}_t}\right) + \ell_t\cdot n_t\cdot w_t\cdot\left(1-\frac{\chi_t}{\bar{\chi}_t}\right)}{\sum_{x\in\mathcal{D}} \ell_{x,t}\cdot n_{x,t}\cdot w_{x,t} + \ell_t\cdot n_t\cdot w_t}.$$

With GE effects, the formula becomes

(A13)

$$\frac{\mathcal{R}_t}{1-\mathcal{R}_t} = \frac{1}{\varepsilon_\ell}\cdot\frac{\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}\cdot\left(1-\frac{\chi_{x,t}}{\bar{\chi}_t}\right)+\ell_t\cdot n_t\cdot w_t\cdot\left(1-\frac{\chi_t}{\bar{\chi}_t}\right)}{\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}+\ell_t\cdot n_t\cdot w_t}$$
$$+\frac{1}{\varepsilon_\ell}\cdot(1-\mathcal{R}_t)\cdot\frac{\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}\cdot\left(\frac{\chi_{x,t}}{\bar{\chi}_t}-1\right)\cdot\frac{d\ln w_{x,t}}{d\mathcal{R}_t}+\ell_t\cdot n_t\cdot w_t\cdot\left(\frac{\chi_t}{\bar{\chi}_t}-1\right)\cdot\frac{d\ln w_t}{d\mathcal{R}_t}}{\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot n_{x,t}\cdot w_{x,t}+\ell_t\cdot n_t\cdot w_t}.$$

When implementing this formula, we can compute the total derivatives $\frac{dw_t}{d\mathcal{R}_t}$ and $\frac{dw_{x,t}}{d\mathcal{R}_t}$ from the system

$$dw_{x,t} = \varepsilon_\ell\cdot\sum_{x'\in\mathcal{D}} f_{y_x,y_{x'}}\cdot\ell_{x',t}\cdot n_{x',t}\left(\frac{dw_{x',t}}{w_{x',t}}-\frac{d\mathcal{R}_t}{1-\mathcal{R}_t}\right)+\varepsilon_\ell\cdot f_{y_x,y}\cdot\ell_t\cdot n_t\cdot\left(\frac{dw_t}{w_t}-\frac{d\mathcal{R}_t}{1-\mathcal{R}_t}\right)$$
$$dw_t = \varepsilon_\ell\cdot\sum_{x'\in\mathcal{D}} f_{y,y_{x'}}\cdot\ell_{x',t}\cdot n_{x',t}\left(\frac{dw_{x',t}}{w_{x',t}}-\frac{d\mathcal{R}_t}{1-\mathcal{R}_t}\right)+\varepsilon_\ell\cdot f_{y,y}\cdot\ell_t\cdot n_t\cdot\left(\frac{dw_t}{w_t}-\frac{d\mathcal{R}_t}{1-\mathcal{R}_t}\right)$$

■

A.3 Proofs for Section 2.5

Proof of Proposition 6. Consider a new path for technology that changes wages by $dw_{x,t}$ and dw_t. The resulting change in welfare is

$$dW_0 = \int_0^\infty\left(\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi_{x,t}\cdot dw_{x,t}+\ell_t\cdot\chi_t\cdot dw_t\right)\cdot dt.$$

Note that changes in reallocation effort are second order. This is because in this case we are varying technology and letting $k_{x,t}$ adjust, which implies that wages are entirely pinned down by technology and independent of reallocation rates.

Differentiating the ideal-price index condition $c^f(\{w_{x,t}\}_{x\in\mathcal{D}}, w_t) = 1$, we obtain

$$s_t\cdot d\ln w_t+\sum_{x\in\mathcal{D}} s_{x,t}\cdot d\ln w_{x,t} = 0.$$

Substituting in the formula for the change in welfare and rearranging terms yields

$$dW_0 = \int_0^\infty\left(\sum_{x\in\mathcal{D}}\left(\ell_{x,t}\cdot\chi_{x,t}\cdot w_{x,t}-\ell_t\cdot\chi_t\cdot w_t\cdot\frac{s_{x,t}}{s_t}\right)\cdot d\ln w_{x,t}\right)\cdot dt.$$

Finally, from $w_{x,t} = 1/A_{x,(1-g)\cdot t}$, we get $d\ln w_{x,t} = \frac{\dot{A}_{x,t}}{A_{x,t}} \cdot t \cdot dg$, which gives the formula in equation (8).

To prove the second part of the Proposition, we show that the government can always change taxes in response to improvements in technology in a way that leaves the economy and welfare unchanged and generates additional revenue. Suppose that $\tau_{x,t}$ is the optimal path for taxes when the productivity of new technology is $A_{x,t}$. Suppose that productivity shifts up to $\tilde{A}_{x,t} \geq A_{x,t}$. Taxes can then be adjusted up to a level $\tilde{\tau}_{x,t}$ defined as

$$\frac{\tilde{A}_{x,t}}{1+\tilde{\tau}_{x,t}} = \frac{A_{x,t}}{1+\tau_{x,t}},$$

and which satisfies $\tilde{\tau}_{x,t} \geq \tau_{x,t}$ while at the same time keeping transfers T_t unchanged. This choice of taxes and transfers ensures that prices, utilities, reallocation rates, and the equilibrium remain unchanged. To conclude the proof we need to show that this policy is feasible, or

$$T_t \leq \sum_{x\in\mathcal{D}} \tilde{\tau}_{x,t} \cdot \frac{k_{x,t}}{\tilde{A}_{x,t}}.$$

This follows from the fact that

$$T_t = \sum_{x\in\mathcal{D}} \tau_{x,t} \cdot \frac{k_{x,t}}{A_{x,t}} = \sum_{x\in\mathcal{D}} \frac{\tau_{x,t}}{1+\tau_{x,t}} \cdot \frac{1+\tilde{\tau}_{x,t}}{\tilde{\tau}_{x,t}} \cdot \tilde{\tau}_{x,t} \cdot \frac{k_{x,t}}{\tilde{A}_{x,t}} \leq \sum_{x\in\mathcal{D}} \tilde{\tau}_{x,t} \cdot \frac{k_{x,t}}{\tilde{A}_{x,t}},$$

as wanted. ■

A.4 Theoretical Extensions

A.4.1 Inequality between and within islands

Our baseline model assumed no ex-ante inequality between or within islands. This is a deliberate choice that helps us isolate the incentives to induce gradual transitions in order to help mitigate the cost of the disruption for affected households. This is in our view what policymakers have in mind when they think of compensating the losers from reforms, globalization, or technological progress. Pre-existing inequities across islands would introduce a very different consideration for taxing trade or new technologies: the possibility of taxing (or subsidizing) a technology because of its "tagging" value as a way to indirectly address pre-existing inequities.

This section provides propositions showing how these considerations affect the optimal taxation problem. We also provide a new rationale for ignoring pre-existing income differences across islands when deciding how to compensate the losers of globalization and technological progress.

Suppose that a household is endowed with $n > 0$ units of labor. We refer to n as the type of the household. The distribution of n in island x has pdf $\varphi_x(n)$, and the distribution of n in undisrupted islands has cdf $\Phi(n)$. This definition implies $\int_n n \cdot d\Phi_x(n) = \ell_{x,0}$ and $\int_n n \cdot d\Phi(n) = \ell_0$.

We make the following assumptions:

- The utility function is homothetic $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$ for some $\gamma > 0$.
- A household of type n receives labor income $\varphi \cdot w_{x,t}$ in island x and a proportional lump-sum rebate of $n \cdot T_t$. It also faces a cost of reallocation $n^{1-\gamma} \cdot \kappa_x(\alpha)$ and is endowed with initial assets $n \cdot a_{0,x}$.
- Labor endowments are transferable across islands. If the household reallocates, it receives labor income $n \cdot w_t$ and a proportional lump-sum rebate $n \cdot T_t$.

These assumptions imply that all households in island x choose paths for consumption and savings that are proportional to each other. In particular $c^n_{x,t_m,t} = n \cdot c_{x,t_m,t}$, $c^n_{x,t} = n \cdot c_{x,t}$, and $\alpha^n_x = \alpha_x$, where $\{c_{x,t_m,t}, c_{x,t}, \alpha_x\}$ are the consumption choices and reallocation effort of a household with one unit of labor. In addition, the households' utility is $\mathcal{U}^n_{x,0} = \mathcal{U}_{x,0} \cdot n^{1-\gamma}$, and their marginal utilities are $\lambda^n_{x,d,t} = \lambda_{x,d,t} \cdot n^{-\gamma}$ and $\lambda^n_{x,r,t} = \lambda_{x,r,t} \cdot n^{-\gamma}$.

Consider a welfare function of the form

$$W_0 = \sum_{x \in \mathcal{D}} \int_n \mathcal{W}\left(\left((1-\gamma) \cdot \mathcal{U}^n_{x,0} \right)^{\frac{1}{1-\gamma}}; n \right) \cdot d\Phi_x(n) + \int_n \mathcal{W}\left(\left((1-\gamma) \cdot \mathcal{U}^h_0 \right)^{\frac{1}{1-\gamma}}; n \right) \cdot d\Phi(n).$$

Relative to what we had in Section 2, this welfare function accounts for heterogeneity in n. Note, however, that we do not require the welfare function to be symmetric, and in particular, we let $\mathcal{W}$ depend on n to capture societal preferences for redistribution across households with different productivities. We also wrote the welfare function in terms of consumption equivalent terms $\left((1-\gamma) \cdot \mathcal{U}^n_{x,0} \right)^{\frac{1}{1-\gamma}}$ and $\left((1-\gamma) \cdot \mathcal{U}^h_0 \right)^{\frac{1}{1-\gamma}}$, but this is done for tractability only.

Proposition A1 *All the results of the paper apply, but with the Pareto weights now re-*

defined as

$$g_x = \int_n \mathcal{W}'\left(n \cdot ((1-\gamma)\cdot \mathcal{U}_{x,0})^{\frac{1}{1-\gamma}}; n\right) \cdot ((1-\gamma)\cdot \mathcal{U}_{x,0})^{\frac{\gamma}{1-\gamma}} \cdot \frac{n \cdot d\Phi_x(n)}{\ell_{x,0}}$$
$$g = \int_n \mathcal{W}'\left(n \cdot ((1-\gamma)\cdot \mathcal{U}_{0})^{\frac{1}{1-\gamma}}; n\right) \cdot ((1-\gamma)\cdot \mathcal{U}_{0})^{\frac{\gamma}{1-\gamma}} \cdot \frac{n \cdot d\Phi(n)}{\ell_{0}}.$$

The proposition illustrates how ex-ante inequality affects optimal policy. Suppose that $\mathcal{W}'(c;n) = c^{-\eta}$ for $\eta \geq \gamma$, so that the welfare function is scale free, symmetric, and concave in individual utilities. Then:

$$g_x = \left(\int_n n^{-\eta} \cdot \frac{n \cdot d\Phi_x(n)}{\ell_{x,0}}\right) \cdot \left(((1-\gamma)\cdot \mathcal{U}_{x,0})^{\frac{1}{1-\gamma}}\right)^{-\eta} \cdot ((1-\gamma)\cdot \mathcal{U}_{x,0})^{\frac{\gamma}{1-\gamma}}$$
$$g = \left(\int_n n^{-\eta} \cdot \frac{n \cdot d\Phi(n)}{\ell_{0}}\right) \cdot \left(((1-\gamma)\cdot \mathcal{U}_{0})^{\frac{1}{1-\gamma}}\right)^{-\eta} \cdot ((1-\gamma)\cdot \mathcal{U}_{0})^{\frac{\gamma}{1-\gamma}}.$$

Ex-ante inequality across households only matters via the terms $\int_n n^{-\eta} \cdot \frac{n \cdot d\Phi_x(n)}{\ell_{x,0}}$ and $\int_n n^{-\eta} \cdot \frac{n \cdot d\Phi(n)}{\ell_{0}}$. These terms are larger for islands x with households that have fewer units of labor on average, introducing a motive for taxing $k_{x,t}$ more aggressively due to its tagging value. On the other hand, within island inequality does not affect optimal taxes conditional on these tagging terms. Optimal taxes are also zero in the long run, since distorting $k_{x,t}$ loses its tagging value as people reallocate away from island x.

The proposition also identifies conditions under which ex-ante inequalities do not interact with the problem of protecting losers. Suppose that $\mathcal{W}$ is of the form

$$\mathcal{W}'(n \cdot c; n) = \mathcal{W}'(c).$$

This captures a situation where the public considers it fair for people with n units of labor to enjoy higher utility and consumption, proportional to their higher human capital. Even if we had lump-sum taxes that could be conditioned on n, we wouldn't make use of them, since inequality along this dimension is deemed fair. In this case

$$g_x = \mathcal{W}'\left(((1-\gamma)\cdot \mathcal{U}_{x,0})^{\frac{1}{1-\gamma}}\right) \cdot ((1-\gamma)\cdot \mathcal{U}_{x,0})^{\frac{\gamma}{1-\gamma}}$$
$$g = \mathcal{W}'\left(((1-\gamma)\cdot \mathcal{U}_{0})^{\frac{1}{1-\gamma}}\right) \cdot ((1-\gamma)\cdot \mathcal{U}_{0})^{\frac{\gamma}{1-\gamma}}$$

and inequality of labor endowments between and within islands is irrelevant for the problem of compensating winners and losers. This offers a rationale for ignoring ex-ante inequalities across (and within) islands when selecting optimal taxes on technologies or trade motivated exclusively by compensating the losers. In particular, we can ignore ex-ante inequalities across islands if they are considered fair.

A.4.2 Congestion externalities

Changes in reallocation effort might also generate congestion, either because of pecuniary externalities (i.e., increasing cost of college enrollment) or congestion (i.e., the search and matching process exhibiting decreasing returns to scale).[39]

Suppose that the reallocation rate is

$$\alpha_x = e_x \cdot (\bar{e}_x \cdot \ell_{x,0})^{\beta-1},$$

where $\bar{e}_x$ is the average reallocation effort e_x chosen by other households in the disrupted island, which comes at an additive utility cost of $\kappa_x(e_x)$. This formulation introduces a congestion externality: households do not internalize that their efforts to reallocate crowd out the reallocation possibilities of others. In this case, households' choice of effort in a symmetric equilibrium within islands satisfies the FOC

$$\mathcal{U}_{x,\alpha} = \kappa_x'(e_x) \cdot (\alpha_x \cdot \ell_{x,0})^{\frac{1-\beta}{\beta}} \qquad \text{where: } e_x = \alpha_x^{\frac{1}{\beta}} \cdot \ell_{x,0}^{\frac{1-\beta}{\beta}}$$

Proposition A2 *Suppose that the reallocation process has decreasing returns to scale of order $\beta \in (0,1]$. A necessary condition for an optimal tax sequence is that*

$$\tau_{x',t} = \sum_{x\in\mathcal{D}} \frac{\ell_{x,t}\cdot w_{x,t}}{m_{x',t}} \cdot \left(\frac{\chi_{x,t}^{\beta end}}{\bar{\chi}_t^{\beta end}} - 1\right) \cdot \left(-\frac{\partial \ln w_{x,t}}{\partial \ln k_{x',t}}\right) + \frac{\ell_t \cdot w_t}{m_{x',t}} \cdot \left(\frac{\chi_t^{\beta end}}{\bar{\chi}_t^{\beta end}} - 1\right) \cdot \left(-\frac{\partial \ln w_t}{\partial \ln k_{x',t}}\right), \tag{A14}$$

where the $\chi^{\beta end}$'s are now given by

$$\begin{aligned}
\chi_{x,t}^{\beta end} =& \chi_{x,t} + \left(\sum_{x''\in\mathcal{D}} \frac{\ell_{x'',0}}{\ell_{x,t}} \cdot \left(\mu_{x''} - \frac{1-\beta}{\beta} \cdot \mathcal{U}_{x'',\alpha}\right) \cdot \varepsilon_{x,x''}\right) \cdot \mathcal{U}_{x,\alpha,d,t},\\
\chi_t^{\beta end} =& \chi_t + \left(\sum_{x''\in\mathcal{D}} \frac{\ell_{x'',0}}{\ell_t} \cdot \left(\mu_{x''} - \frac{1-\beta}{\beta} \cdot \mathcal{U}_{x'',\alpha}\right) \cdot \varepsilon_{x,x''}\right) \cdot \mathcal{U}_{x,\alpha,r,t}.
\end{aligned}$$

Proof of Proposition A2. Recall that the cost of reallocating at a rate α_x is $\kappa(e_x)$, which can be written as

$$\kappa\left(\alpha_x^{\frac{1}{\beta}} \cdot \ell_{x,0}^{\frac{1-\beta}{\beta}}\right).$$

[39] Pecuniary externalities do not necessarily involve an inefficiency. However, they matter from a welfare point of view because they reduce disposable income for affected households.

We can compute the welfare gains from the reform that changes $k_{x',t}$ by $dk_{x,t}$ and leaves all other $k_{x,s}$ unchanged as

$$\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi_{x,t}\cdot\left(d_k w_{x,t}+d_k T_t\right)\cdot dt+\ell_t\cdot\chi_t\cdot\left(d_k w_t+d_k T_t\right)\cdot dt$$
$$+\sum_{x\in\mathcal{D}}\mu_x\cdot\ell_{x,0}\cdot d\alpha_x+\sum_{x\in\mathcal{D}}\ell_{x,0}\cdot\left(U_{x,\alpha}-\kappa'(e_x)\cdot\frac{1}{\beta}\cdot\left(e_x\cdot\ell_{x,0}\right)^{\frac{1-\beta}{\beta}}\right)\cdot d\alpha_x=0.$$

Using households' first order condition for reallocation effort, we can rewrite this as

$$\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi_{x,t}\cdot\left(d_k w_{x,t}+d_k T_t\right)\cdot dt+\ell_t\cdot\chi_t\cdot\left(d_k w_t+d_k T_t\right)\cdot dt$$
$$+\sum_{x\in\mathcal{D}}\mu_x\cdot\ell_{x,0}\cdot d\alpha_x+\sum_{x\in\mathcal{D}}\ell_{x,0}\cdot\left(U_{x,\alpha}-\frac{1}{\beta}\cdot\mathcal{U}_{x,\alpha}\right)\cdot d\alpha_x=0,$$

or

$$\sum_{x\in\mathcal{D}}\ell_{x,t}\cdot\chi_{x,t}\cdot\left(d_k w_{x,t}+d_k T_t\right)\cdot dt+\ell_t\cdot\chi_t\cdot\left(d_k w_t+d_k T_t\right)\cdot dt$$
$$+\sum_{x\in\mathcal{D}}\mu_x\cdot\ell_{x,0}\cdot d\alpha_x-\sum_{x\in\mathcal{D}}\ell_{x,0}\cdot\frac{1-\beta}{\beta}\cdot U_{x,\alpha}\cdot d\alpha_x=0,$$

This shows that the formulas for optimal taxes are the same as above but with $\mu_x-\frac{1-\beta}{\beta}\cdot U_{x,\alpha}$ in place of μ_x, which accounts for the congestion externality.

Note that $\varepsilon_{x,x''}$ has the same definition as above, but solves a lightly different equation in terms of primitives. In particular, equation (A3) becomes

(A15)
$$\kappa''_\beta(\alpha_x)\cdot d\alpha_x=\sum_{x''}\theta_{x,x''}\cdot d\alpha_{x''}+\mathcal{U}_{x,\alpha,d,t}\cdot\left(d_k w_{x,t}+d_k T_t\right)\cdot dt+\mathcal{U}_{x,\alpha,r,t}\cdot\left(d_k w_t+d_k T_t\right)\cdot dt,$$

where the function κ_β satisfies

$$\kappa'_\beta(\alpha_x)=\kappa'\left(\alpha_x^{\frac{1}{\beta}}\cdot\ell_{x,0}^{\frac{1-\beta}{\beta}}\right)\cdot\left(\alpha_x\cdot\ell_{x,0}\right)^{\frac{1-\beta}{\beta}},$$

so that it captures the extra curvature introduced by decreasing returns to reallocation. ■

A.4.3 Retraining subsidies

PROPOSITION A3 *Suppose the planner has other policy tools that implement the optimal social level of reallocation. A necessary condition for an optimal tax sequence is that:*

$$\tau_{x',t} = \sum_{x\in\mathcal{D}} \frac{\ell_{x,t}\cdot w_{x,t}}{m_{x',t}}\cdot\left(\frac{\chi_{x,t}}{\bar{\chi}_t}-1\right)\cdot\left(-\frac{\partial\ln w_{x,t}}{\partial\ln k_{x',t}}\right)+\frac{\ell_t\cdot w_t}{m_{x',t}}\cdot\left(\frac{\chi_t}{\bar{\chi}_t}-1\right)\cdot\left(-\frac{\partial\ln w_t}{\partial\ln k_{x',t}}\right), \tag{A16}$$

where the multipliers on the right-hand side are now evaluated along an equilibrium with the socially optimal level of α_x.

Proof of Proposition A3. Optimal reallocation effort maximizes social welfare. The envelope theorem implies that the effect of any reform on welfare is equal to the direct effect holding α_x constant, which leads to the same optimality condition as in Proposition 3. ■

A.5 Calibration and Details of Numerical Algorithms

A.5.1 Calibration Details, China Shock

Calibrating π: Industry prices are initially given by $P_i = 1$. Following the disruption, we get a price index

$$P_{i,t_f} = c_i(W_t, \exp(-\pi)),$$

for some cost function c_i with $c_i(1,1) = 1$. Assuming that π is small, we can log-linearize this equation around $(1,1)$ as

$$\ln P_{i,t_f} \approx \text{share domestic production}_{i,t_f}\cdot\ln W_t - \text{share Chinese production}_{i,t_f}\cdot\ln\pi.$$

This implies

$$\ln P_{i,t_f} =\approx \text{share domestic production}_{i,t_f}\cdot(\ln W_t + \ln\pi) - \ln\pi.$$

Let $s = \max\{\text{share Chinese production}_{i,t_f}\}$ and suppose that s is small, as is the case in the data. Then

$$\ln P_{i,t_f} =\approx \ln\text{share domestic production}_{i,t_f}\cdot\ln\pi - \ln\pi.$$

This shows that the regression in Bai and Stumpner (2019) across industries identifies $\ln\pi$.

Pre-existing trade: In the applications of our framework to the China Shock and Colombia's trade liberalization, we have to deal with the fact that there was some pre-existing trade.

For the China Shock, we handle pre-existing trade by assuming that there is a mass $\nu_{p(i)}$ of islands associated with industry i that were already replaced by Chinese imports and hosted no workers by 1991. We normalize the cost of Chinese imports in these islands to 1, which implies that the cost function associated with (11) becomes

$$c_u^f(\{w_x\}, w) = \left(\nu_p + \nu \cdot w^{1-\sigma} + \sum_{x \in \mathcal{D}} \nu_x \cdot w_x^{1-\sigma}\right)^{\frac{1}{1-\sigma}},$$

where $\nu_p = \sum_i \nu_{p(i)}$. The normalization $\bar{w} = 1$ in status quo then requires $\nu_p + \nu + \sum_{x \in \mathcal{D}} \nu_x = 1$. In our calibration, we set $\nu_p = 2.5\%$—the share of imports in GDP before the China Shock.

We assume that the China Shock is driven by advances in the productivity of Chinese imports at other islands, and not by cost reductions of established products. These assumptions imply that the status-quo level of imports in industry i is

$$\frac{m_{i,t_0}}{y_{i,t_0}} = \nu_{p(i)};$$

while imports in industry i at time t after the China Shock are given by

$$\frac{m_{i,t}}{y_{i,t}} = \nu_{p(i)} + \frac{1}{y_t \cdot A_{x,t}} \cdot \left(\ell_t \cdot \frac{c_x^f}{c_w^f} - \ell_{x,t}\right),$$

where x is defined as the island associated with industry i (i.e., the one for which $i(x) = i$). In this expression, the first term accounts for imports at islands with pre-existing trade and the second term accounting for imports in new islands. The change in normalized import shares at time t is then equal to

$$\text{(A17)} \quad \text{Change in normalized import share}_{i(x),t} = \frac{1}{\Omega_i} \cdot \frac{1}{y_t \cdot A_{x,t}} \cdot \left(\ell_t \cdot \frac{c_x^f}{c_w^f} - \ell_{x,t}\right) \text{ for } x \in \mathcal{D}.$$

Equation (A17) provides a system of equation across industries that we use to calibrate $\nu_{i(x)}$ and $s_{i(x)}$ in a first step to match the change in normalized imports by 2007 (recall that $A_{x,t_f} = \exp(\pi)$ at this point), and then to calibrate a path for $A_{x,t}$ in a second step, as described in the main text.

For Colombia's trade liberalization, we assume that a mass $\nu_{p(i)}$ of segments were already produced internationally and hosted no workers by 1989. In addition, we assume this segments were not protected by 1989, and experienced no subsequent decline in tariffs after the 1990 trade liberalization. Under these assumptions, we have that the status-quo level of imports in industry i is

$$\frac{m_{i,t_0}}{y_{i,t_0}} = \nu_{p(i)};$$

while imports in industry i at time t after the liberalization are

$$\frac{m_{i,t}}{y_{i,t}} = \nu_{p(i)} + \frac{1+\tau_{x,t}}{y_t \cdot A_{x,t}} \cdot \left(\ell_t \cdot \frac{c_x^f}{c_w^f} - \ell_{x,t} \right),$$

where x is defined as the island associated with industry i (i.e., the one for which $i(x) = i$). The change in normalized import shares at time t is then equal to

$$\text{Change in normalized import share}_{i(x),t} = \frac{1}{\Omega_i} \cdot \frac{1+\tau_{x,t}}{y_t \cdot A_{x,t}} \cdot \left(\ell_t \cdot \frac{c_x^f}{c_w^f} - \ell_{x,t} \right) \text{ for } x \in \mathcal{D}. \tag{A18}$$

Equation (A18) provides a system of equation across industries that we use to calibrate $\nu_{i(x)}$ and $s_{i(x)}$ to match the increase in normalized import shares between 1989 and 2002.

A.5.2 Implementing the formulas in equations (3) and (4):

This subsection describes the numerical procedure used to compute optimal taxes.

Exogenous reallocation effort. We compute optimal taxes with exogenous reallocation effort as follows:

1. Start with $\tau_{x,t}^{(0)} = 0$ (Laissez Faire).

2. Compute equilibrium objects for $\tau_{x,t}^{(n)}$, identified with the superscript (n) below.

3. Use equation (3) to update optimal taxes as

$$\tau_{x',t}^{(n+1)} = \sum_{x \in \mathcal{D}} \frac{\ell_{x,t} \cdot w_{x,t}^{(n)}}{m_{x',t}^{(n)}} \cdot \left(\frac{\chi_{x,t}^{(n)}}{\bar{\chi}_t^{(n)}} - 1 \right) \cdot \left(- \frac{\partial \ln w_{x,t}^{(n)}}{\partial \ln k_{x',t}^{(n)}} \right) + \frac{\ell_t \cdot w_t^{(n)}}{m_{x',t}^{(n)}} \cdot \left(\frac{\chi_t^{(n)}}{\bar{\chi}_t^{(n)}} - 1 \right) \cdot \left(- \frac{\partial \ln w_t^{(n)}}{\partial \ln k_{x',t}^{(n)}} \right),$$

4. Repeat steps 2–3 until convergence.

Endogenous reallocation effort. We compute optimal taxes with exogenous reallocation effort as follows:

1. Start with $\tau_{x,t}^{(0)} = 0$ (Laissez Faire) and the observed rate of reallocation $\alpha_x^{(0)}$.

2. Compute equilibrium objects for $\tau_{x,t}^{(n)}$ and $\alpha_x^{(0)}$, identified with the superscript (n) below.

3. Use equation (3) to update optimal taxes as

$$\tau_{x',t}^{(n+1)} = \sum_{x\in\mathcal{D}} \frac{\ell_{x,t}\cdot w_{x,t}^{(n)}}{m_{x',t}^{(n)}} \cdot \left(\frac{\chi_{x,t}^{\text{end},(n)}}{\bar{\chi}_t^{\text{end},(n)}} - 1\right) \cdot \left(-\frac{\partial \ln w_{x,t}^{(n)}}{\partial \ln k_{x',t}^{(n)}}\right) + \frac{\ell_t \cdot w_t^{(n)}}{m_{x',t}^{(n)}} \cdot \left(\frac{\chi_t^{\text{end},(n)}}{\bar{\chi}_t^{\text{end},(n)}} - 1\right) \cdot \left(-\frac{\partial \ln w_t^{(n)}}{\partial \ln k_{x',t}^{(n)}}\right),$$

4. Update the reallocation rate using

$$\alpha_x^{(n+1)} = \alpha_x^{(n)} + \Delta\alpha_x^{(n)},$$

   where $\Delta\alpha_x^{(n)}$ is given by

$$\Delta\alpha_x^{(n)} = \sum_{x''} \varepsilon_{x,x''} \cdot \int_0^\infty \left(\mathcal{U}_{x'',\alpha,d,t}^{(n)} \cdot (\Delta^{(n)} w_{x'',t} + \Delta^{(n)} T_t) + \mathcal{U}_{x'',\alpha,r,t}^{(n)} \cdot (\Delta^{(N)} w_t + \Delta^{(n)} T_t)\right) \cdot dt,$$

   and $\Delta^{(n)} w_{x'',t}$, $\Delta^{(n)} w_t$, $\Delta^{(n)} T_t$ is the change in wages and tax revenue generated by the update in taxes from iteration n to $n+1$.

5. Repeat steps 2–4 until convergence.

This procedure only requires us to specify values for $\varepsilon_{x,x''}$ and solve for the optimal tax and the equilibrium path without having to specify the κ function. This comes at the cost of assuming that the elasticities $\varepsilon_{x,x''}$ remain roughly unchanged for the variations in taxes considered. It also ignores the effect of changes in household utility on the multipliers g_x, which is second order due to the envelope theorem, but could be non-negligible for large changes in reallocation effort α_x.

As an alternative, we experimented with the following procedure, which requires parametrizing the κ_x function:

1. Start with $\tau_{x,t}^{(0)} = 0$ (Laissez Faire) and the observed rate of reallocation $\alpha_x^{(0)}$.

2. Compute equilibrium objects for $\tau_{x,t}^{(n)}$ and $\alpha_x^{(0)}$, identified with the superscript (n) below.

3. Compute $\varepsilon^{(n)}_{x,x''}$ by solving the system of equations in (A2).

4. Use equation (3) to update optimal taxes as

$$\tau^{(n+1)}_{x',t} = \sum_{x\in\mathcal{D}} \frac{\ell_{x,t}\cdot w^{(n)}_{x,t}}{m^{(n)}_{x',t}} \cdot \left(\frac{\chi^{\text{end},(n)}_{x,t}}{\bar{\chi}^{\text{end},(n)}_{t}} - 1\right) \cdot \left(-\frac{\partial \ln w^{(n)}_{x,t}}{\partial \ln k^{(n)}_{x',t}}\right) + \frac{\ell_{t}\cdot w^{(n)}_{t}}{m^{(n)}_{x',t}} \cdot \left(\frac{\chi^{\text{end},(n)}_{t}}{\bar{\chi}^{\text{end},(n)}_{t}} - 1\right) \cdot \left(-\frac{\partial \ln w^{(n)}_{t}}{\partial \ln k^{(n)}_{x',t}}\right),$$

   using the values of $\varepsilon^{(n)}_{x,x''}$ to compute the χ's.

5. Update the reallocation rate using

$$\kappa'(\alpha^{(n+1)}_{x}) = \mathcal{U}^{(n)}_{x,\alpha}.$$

6. Repeat steps 2–5 until convergence.

A.5.3 Details of the savings problem with no risk sharing:

As explained in the text, households problem can be summarized by the following system of HJB equations

$$\rho v_x(a,t;\alpha) - \dot{v}_x(a,t;\alpha) = \max_c u(c) + \partial_a v_x(a,t;\alpha)\cdot(ra + w_{x,t} - c) + \alpha_x\cdot(v(a,t) - v_x(a,t;\alpha)),$$
$$\rho v(a,t) - \dot{v}(a,t) = \max_c u(c) + \partial_a v(a,t)\cdot(ra + w_t - c).$$

Here, $v_x(a,t;\alpha)$ is the value function of households in disrupted islands at time t with assets a when they exert reallocation effort α, and $v(a,t)$ is the value function of households in undisrupted islands with assets a.

Let $h_{x,t} = \int_t^\infty e^{-(s-t)r}\cdot w_{x,s}ds$ and $h_t = a_t + \int_t^\infty e^{-(s-t)r}\cdot w_s ds$. We can rewrite these HJB equations using $z = a + h$—effective wealth—as the state variable:

$$\rho v_x(z,t;\alpha) - \dot{v}_x(z,t;\alpha) = \max_c u(c) + \partial_z v_x(z,t;\alpha)\cdot(rz - c) + \alpha_x\cdot(v(z + h_t - h_{x,t}) - v_x(z,t;\alpha)),$$
$$\rho v(z) = \max_c u(c) + \partial_z v(z)\cdot(rz - c).$$

Note that the HJB equation for $v(z)$ is now stationary, since interest rates are constant. For $u(c) = c^{1-\gamma}/(1-\gamma)$, we can solve analytically for $v(z)$ as

$$v(z) = \left[r - \frac{1}{\gamma}(r-\rho)\right]^{-\gamma}\cdot\frac{z^{1-\gamma}}{1-\gamma}.$$

Moreover, policy functions in the undisrupted island are given by

$$c_t = \left[r - \frac{1}{\gamma}(r-\rho)\right] \cdot z_t, \dot{z}_t = \qquad \frac{1}{\gamma}(r-\rho) \cdot z_t.$$

This implies

$$\lambda_{x,r,t} = \frac{1}{1-P_{x,t}} \cdot \int_0^t e^{-\rho t} \cdot \alpha_x \cdot e^{-\alpha_x t_r} \cdot \left(\left[r - \frac{1}{\gamma}(r-\rho)\right] \cdot (z_{x,t_r} + h_{t_r} - h_{x,t_r}) \cdot e^{\frac{1}{\gamma}(r-\rho)(t-t_r)}\right)^{-\gamma} \cdot dt_r,$$

where $z_{x,t}$ denotes the effective wealth held by households in disrupted islands at time t. This expression uses the fact that

$$c_{x,t_r,t} = \left[r - \frac{1}{\gamma}(r-\rho)\right] \cdot (z_{x,t_r} + h_{t_r} - h_{x,t_r}) \cdot e^{\frac{1}{\gamma}(r-\rho)(t-t_r)}.$$

To characterize $z_{x,t}$ we solve the HJB equation for $v_x(z, t; \alpha)$ numerically using the finite-differences method described in Achdou et al. (2021). This method characterizes the common path of consumption $c_{x,t}$ and assets $z_{x,t}$ for households in disrupted islands starting from $z_{x,0} = a_{x,0} + h_{x,0}$. From this method, we also obtain

$$\lambda_{x,d,t} = e^{-\rho t} \cdot c_{x,t}^{-\gamma}.$$

Figure A1 plots typical path for consumption $c_{x,t}$ and assets $z_{x,t}$ starting from $z_{x,0} = h_{x,0} = 1$ in an economy where $h_t - h_{x,t}$ is positive and rises from 0.3 to 0.5 over time. For this examples, we consider a baseline scenario with $\alpha_x = 5\%$, $r = \rho = 5\%$, and $\gamma = 2$ and report variants.

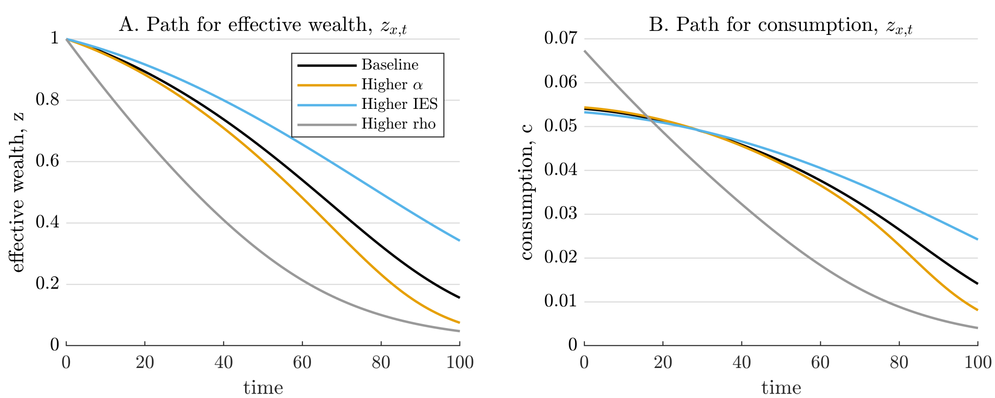

FIGURE A1: CONSUMPTION AND WEALTH PATH IN DISRUPTED ISLANDS. The figure reports examples of the optimal path for effective wealth and consumption in disrupted islands when households can borrow but face uncertainty regarding when they will reallocate. These paths are obtained numerically using the finite-differences method described in Achdou et al. (2021).