Globalization, Inequality and Welfare*

Pol Antràs
pantras@fas.harvard.edu

Alonso de Gortari
adegortaribriseno@fas.harvard.edu

Oleg Itskhoki
itskhoki@princeton.edu

September 19, 2016

Abstract

This paper studies the welfare implications of trade opening in a world in which trade raises aggregate income but also increases income inequality, and in which redistribution needs to occur via a distortionary income tax-transfer system. We provide tools to characterize and quantify the effects of trade opening on the distribution of *disposable* income (after redistribution). We propose two adjustments to standard measures of the welfare gains from trade: a 'welfarist' correction inspired by the Atkinson (1970) index of inequality, and a 'costly-redistribution' correction capturing the efficiency costs associated with the behavioral responses of agents to trade-induced shifts across marginal tax rates. We calibrate our model to the United States over the period 1979-2007 using data on the distribution of adjusted gross income in public samples of IRS tax returns, as well as CBO information on the tax liabilities and transfers received by agents at different percentiles of the U.S. income distribution. Our quantitative results suggest that both corrections are nonnegligible: trade-induced increases in inequality of disposable income erode about 20% of the gains from trade, while the gains from trade would be about 15% larger if redistribution was carried out via non-distortionary means.

*Harvard University, Harvard University, and Princeton University. We thank Nathan Hendren and Esteban Rossi-Hansberg for fruitful conversations during the early stages of this project. This paper was previously circulated under the title 'Inequality, Costly Redistribution and Welfare in an Open Economy'. We are indebted to Xiang Ding and Teresa Fort for their help obtaining a key moment in our calibration from U.S. Census data, and to Daniel Feenberg for his assistance with the NBER-IRS data. We are also grateful to seminar audiences at LSE, Barcelona Summer Forum, Degit XX in Geneva, SED in Warsaw, AEEFI in San Sebastián, Penn State, ITSG in Rome, IMF, Princeton, Stanford, UBC, National University of Singapore, Lausanne, NOITS in Oslo, Bocconi, Seoul National University, Bayreuth, Harvard, and the Federal Reserve Board for their suggestions.

1 Introduction

Two of the most salient phenomena in the world economy in recent years have been a rapid increase in the extent to which economies have become interconnected and a significant rise in income inequality in many countries. For instance, during the period 1979-2007, the U.S. trade share (defined as the average of exports and imports divided by U.S. gross output) increased from a value of 4.9% to 7.7%, while the Gini coefficient associated with the distribution of U.S. market income grew dramatically from a level of 0.48 all the way to 0.59. Furthermore, as is clear from Figure 1, trade integration and inequality grew very much in parallel even at fairly high frequencies. The extent to which these two phenomena are causally related has been the subject of intense academic debates, but it is by now a widely accepted view that trade integration has been a significant contributor to increased wage and income inequality in the U.S. and many other industrialized countries.[1] The picture emerging from developing countries also points to the importance of trade-induced inequality. Goldberg and Pavcnik (2007) summarize a body of literature studying the consequences of trade liberalization across a number of developing countries after 1970s, with the bulk of episodes triggering significant increases in inequality.

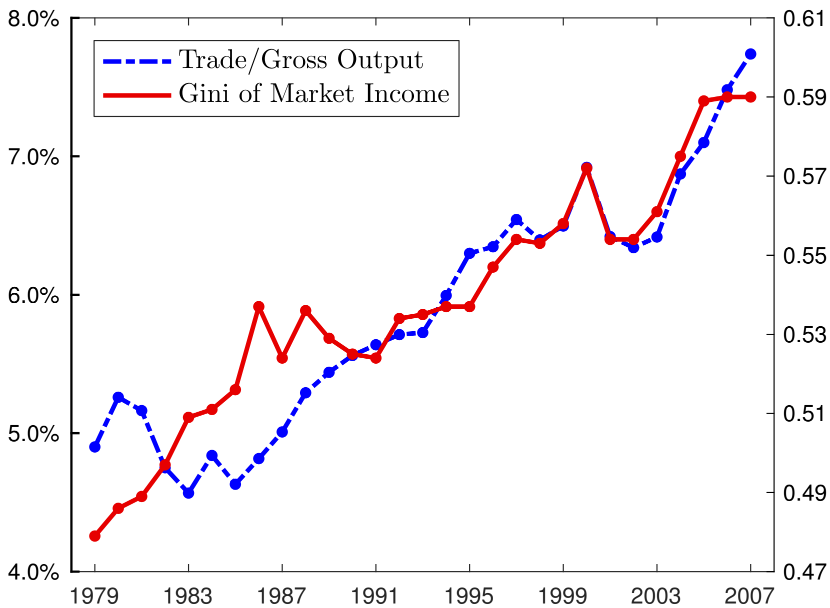

Figure 1: Trade Integration and Income Inequality: United States (1979-2007)

Despite these recent trends, the standard approach to demonstrating and quantifying the welfare gains from trade largely ignores the implications of trade-induced inequality. The paradigm used to evaluate the social welfare consequences of trade integration is the Kaldor-Hicks compensation principle (Kaldor, 1939; Hicks, 1939). This approach begins by computing the compensating (or equivalent) variation of a policy change at the individual level, and then aggregates

[1] Feenstra and Hanson (1999), for instance, estimate that outsourcing alone could account for as much as 40% of the increase in the U.S. skill premium in the 1980s. Other studies, summarized in Krugman (2008), arrive at more conservative estimates suggesting that trade accounted for about 15-20% of the increase in income inequality.

this money metric across agents. The celebrated 'gains from trade' result demonstrates that, in competitive environments, when moving from autarky to any form of trade integration, the losers can always be compensated and there is some surplus to potentially turn this liberalization into a Pareto improvement. A key advantage of the Kaldor-Hicks criterion as a tool for policy evaluation is that it circumvents the need to base policy recommendations on interpersonal comparisons of utility, thus extricating economists' prescriptions from their own moral convictions (cf., Robbins, 1932).

As influential as the Kaldor-Hicks compensation principle has proven to be, it has two basic shortcomings. First, the fact that there is the *potential* to compensate those that are hurt from a particular policy does not imply that these losers will be compensated *in practice*. If one knew that the redistribution or compensation necessary for a policy to generate Pareto gains would not happen or would not be complete, shouldn't the evaluation of such a policy take this fact into account? Second, the simple aggregation of individual compensating or equivalent variations in the Kaldor-Hicks criterion implicitly assumes the existence of nondistortionary means to redistribute part of the gains from the policy to those that do not directly benefit from it. In reality, compensation often takes place through a tax and transfer system embodying nontrivial deadweight losses, so it seems reasonable to build this characteristic of redistribution into measures of the social welfare effects of a policy.

In this paper, we study the welfare implications of trade opening in a world in which international trade affects the shape (and not just the mean) of the income distribution, and in which redistribution policies need to occur via a distortionary income tax-transfer system. In this environment, we provide tools to characterize and quantify the *actual* amount of compensation that will take place following trade opening, as well as the *efficiency costs* of undertaking such redistribution. More specifically, we propose two types of adjustments to standard measures of the welfare gains from trade. On the one hand, we develop a 'welfarist' correction which captures the negative impact that an increase in inequality in the distribution of disposable incomes has on the welfare of an inequality-averse social planner. This first adjustment is tightly related to the Atkinson (1970) index of inequality, which has been rarely applied in trade contexts.[2] On the other hand, we derive a 'costly-redistribution' correction which captures the behavioral responses of agents to *trade-induced* shifts across marginal tax rates. This second adjustment builds on the voluminous public finance literature on the efficiency costs of income taxation, and is especially related to the structural work of Benabou (2002), although our approach is generalized to apply to income distributions other than the lognormal one, and also to models with an extensive margin response to taxation.

We begin our analysis in section 2 within a fairly general environment that illustrates the rationale for these two corrections when evaluating *any* policy (not just trade liberalization) that has the potential to affect the shape of the income distribution beyond its mean. In this environment, we derive explicit formulas for these adjustments in terms of specific moments of the income distribution, the level of progressivity of the tax-transfer system, the degree of

[2] Three very recent exceptions are the ongoing projects by Rodriguez-Clare, Galle, and Yi (2015), Porto (2015) and Carrère, Grujovic, and Robert-Nicoud (2015).

inequality aversion of the social planner, and the elasticity of taxable income to changes in marginal tax rates.

Our environment in section 2 is silent on the primitive determinants of the income-generation process or on the precise mechanism that leads to a positive elasticity of income to changes in marginal taxes. In section 3, we develop a microfounded simple general equilibrium framework that illustrates how the ability of individuals and their labor supply decisions translate in equilibrium earnings and welfare levels given the tax system in place. When solving for the closed-economy equilibrium of the model, we are able to decompose changes in welfare into three terms: (i) changes in the welfare of a hypothetical 'Kaldor-Hicks' economy with access to costless redistribution and no inequality aversion, (ii) changes in the welfarist correction, and (iii) changes in the costly-redistribution correction.

The economic environment we develop builds on Itskhoki (2008), and is inspired by the canonical optimal taxation framework of Mirrlees (1971) and the workhorse model of trade of Melitz (2003). Agents in our economy are workers each producing a distinct task associated with the production of a final good. Unobservable heterogeneity in productivity across agents generates income inequality, which an inequality-averse social planner may try to moderate via a progressive system of income taxation. The two key departures from the classic Mirlees framework is that we allow for imperfect substitutability in the task services provided by different workers and that we restrict attention to a specific form of nonlinear taxation that, consistently with U.S. data, implies a log-linear relationship between income levels before and after taxes and transfers (see Heathcote, Storesletten, and Violante, 2014). Imperfect substitutability is not essential for our closed-economy results but is the source of the welfare gains from trade later in the paper.[3]

Before moving to this open-economy environment, in section 4 we provide a brief calibration of the closed-economy model that decomposes the evolution of social welfare in the U.S. over the period 1979-2007 in terms of the welfarist and costly-redistribution corrections and the welfare of the hypothetical 'Kaldor-Hicks' economy. We calibrate our model using data on the distribution of adjusted gross income in public samples of IRS tax returns, as well as CBO information on the tax liabilities and transfers received by agents at different points of the U.S. income distribution. Our calibration reveals a very significant decline in the degree of tax progressivity over this period despite the concomitant increase in 'primitive' income inequality. This naturally resulted in an exacerbated increase in inequality in the distribution of *disposable* income. As a consequence, even for modest degrees of inequality aversion, the implied social welfare gains are significantly lower than the average real income gains recorded over this period.[4]

[3] Imperfect substitutability between different types of labor in the Mirrlees model was studied by Feldstein (1973) and Stiglitz (1982) in a two-class economy, and more recently by Sachs, Tsyvinski, and Werquin (2016) in a more general environment.

[4] Throughout the paper we equate consumption with disposable income. An active literature in macroeconomics has discussed the extent to which consumption inequality has tracked income inequality in recent decades. Using data from the U.S. Consumer Expenditure Survey, Krueger and Perri (2006) initially found that consumption inequality had grown much less than income inequality between 1980 and 2004. Nevertheless, Attanasio, Hurst, and Pistaferri (2012) and Aguiar and Bils (2015) have shown that when properly addressing measurement error biases, consumption inequality within the U.S. appears to have increased between 1980 and 2010 by nearly the

We also use our simple calibration to shed light on the growth in average real income that the U.S. would have attained if the progressivity of the U.S. tax system had been kept constant at its 1979 level, or if tax progressivity had increased to avoid the observed rise in income inequality. We find that real income growth in those counterfactual scenarios would have been markedly lower than the 1.31% annual increase observed in the data. For instance, for the case of a degree of inequality aversion equal to 1 and an elasticity of taxable income equal to 0.5, we find that real income would have grown at an average annual rate of 0.85% if tax progressivity had remained constant, and at an even lower annual rate of 0.40% if progressivity had increased to keep income inequality unchanged over the period.

Armed with this suggestive evidence of the empirical relevance of our two key inequality corrections, in section 5 we move to an open-economy environment, which is a direct extension of the closed-economy framework in section 3. In particular, our assumed imperfect substitutability of the tasks performed by different workers worldwide results in welfare gains from trade integration associated with final output being produced with a wider range of differentiated tasks. These love-for-variety gains from trade are thus analogous to those in Krugman (1980) or Ethier (1982). In order to generate nontrivial effects of trade on the income distribution, we follow Melitz (2003) and introduce fixed costs of exporting, which allow only the most productive agents to participate in international trade. Consequently, trade disproportionately benefits the most productive agents in society, leading to greater income inequality in a trading equilibrium than under autarky. The progressivity of the tax system attenuates the rise in inequality following trade liberalization, but unless tax progressivity increases with trade, the distribution of disposable income will necessarily become more unequal with trade, thus leading to a higher welfarist correction than under autarky. Furthermore, selection into exporting generates an extensive margin of trade, which is sensitive to national redistribution policies and contributes to the overall efficiency costs of taxation. We thus find that our 'costly-redistribution' correction is also generally exacerbated by a process of trade integration.

In section 6, we calibrate our open-economy model with the same IRS tax returns data employed in section 4, together with measures of trade exposure to calibrate the key trade frictions parameters of the model. We then perform counterfactuals to gauge the effects of trade on aggregate income and inequality. Our quantitative results suggest that our two suggested welfare corrections are nonnegligible. Under our preferred parametrization, trade-induced increases in disposable income inequality erode about 20% of the U.S. gains from trade, while gains from trade would have been about 15% larger if redistribution had been carried out via non-distortionary means. The size of the two corrections also appears to be fairly insensitive to alternative parametrizations of the model.

Our model of the effects of trade on the income distribution is highly stylized and abstracts from many features that have been emphasized in past and more recent research on trade and labor markets. For many years, the Heckscher-Ohlin (HO) model, and in particular its Stolper-Samuelson theorem, provided the key conceptual framework used to analyze the links between

same amount as income inequality. See also Jones and Klenow (2016) for welfare measures that incorporate the role of consumption, leisure, mortality, and inequality.

trade and wage inequality. Nevertheless, the empirical limitations of this framework have become apparent in recent years. As mentioned above, sharp increases in inequality happened not only in rich but also in unskilled-labor abundant developing countries, a phenomenon at odds with the predictions of the HO model.[5] In addition, the contribution of the residual component of wage inequality within groups of workers with similar observable characteristics appears to be at least as important as the growing skill premium across groups, which is the only channel captured by the HO model.[6] Finally, contrary to the main mechanism of adjustment in the HO model, the reallocation within sectors appears to be more important than across sectors for both adjustment to trade and inequality dynamics.[7]

For these reasons, recent work has explored alternative models featuring richer interactions between labor markets and trade liberalization. One branch of this literature has explored the role of search frictions and other types of labor-market imperfections (see, for instance, Helpman, Itskhoki, and Redding, 2010, Egger and Kreickemeier, 2009a or Amiti and Davis, 2012, among many others), while a second branch has focused on the role of sorting of heterogeneous workers into firms or technologies (see, for instance, Yeaple, 2005, Costinot and Vogel, 2010 or Sampson, 2014), or the matching of heterogeneous workers into production teams (see Antràs, Garicano, and Rossi-Hansberg, 2006). Our international trade model is more parsimonious than those developed in this recent research, yet the mechanism through which it generates trade-induced inequality is the same. The key distinction of our stylized model, critical for our analysis, is that it allows us to incorporate behavioral responses to taxation, not featured in previous work. Another important advantage for our purposes is that, despite being stylized, our model is readily amenable to the calibration of the full income distribution and the quantification of the counterfactual inequality effects from a trade liberalization. An open question for future research is the extent to which the inequality corrections arising from our framework are similar in magnitude to those one would obtain in richer frameworks.[8]

Within the international trade field, our paper is also related to previous work studying the redistribution of the gains from trade. Following Dixit and Norman (1980, 1986), this strand of the literature has mainly focused on the possibility of compensating the losers from trade through a variety of tax instruments. Dixit and Norman themselves focused on the sufficiency of commodity and factor taxation for ensuring Pareto gains from trade, while Spector (2001) and Naito (2006) showed how Mirrlees-type incentive constraints could undermine differential factor taxation, thereby opening the door for the possibility that trade could lead to welfare losses by

[5]A related observation is that the movements in relative prices of skilled to unskilled goods, which are at the core of the Stolper-Samuelson mechanism, tended to be small (e.g., see Lawrence and Slaughter, 1993).

[6]For example, see Autor, Katz, and Kearney (2008) for the evidence for US and Attanasio, Goldberg, and Pavcnik (2004) for the evidence for a developing country (Colombia).

[7]For example, Faggio, Salvanes, and Van Reenen (2007) show that most of the increase in wage inequality in the U.K. happened within industries, while Levinsohn (1999) shows the relative importance of within-industry reallocation in response to trade liberalization in Chile. See also Burstein and Vogel (2016).

[8]For instance, it would be interesting to explore the robustness of our results to an environment with nonhomothetic preferences. Whether trade integration has increased or decreased income inequality through its effect on the price index relevant to individuals with different levels of income is an open empirical question (see Broda and Romalis, 2008, Fajgelbaum and Khandelwal, 2016, or Jaravel, 2016).

hampering redistribution.[9] Relative to this body of work, our goal is to instead characterize and quantify the *actual* efficiency costs of redistribution given the observed features of the system used to carry out such compensation in the real world.[10] In that sense, our focus on the income tax-transfer system as the vehicle for redistribution is motivated by the small scale and limited relevance of more direct means of compensation, such as trade adjustment assistance programs. For instance, in their influential recent study on the U.S. labor-market implications of the rise of Chinese import competition, Autor, Dorn, and Hanson (2013) find that the estimated dollar increase in per capita Social Security Disability Insurance (SSDI) payments following trade-induced job displacements is more than thirty times as large as the estimated dollar increase in Trade Adjustment Assistance (TAA) payments.

Finally, our welfarist and costly redistribution corrections are not only related to the contributions of Kaldor (1939), Hicks (1939), Atkinson (1970), and Benabou (2002), but they also connect to a large body of related work. The welfarist approach to policy evaluation originates in the pioneering work of Bergson (1938) and Samuelson (1948), and has constituted an important paradigm in the optimal policy literature since the seminal work of Diamond and Mirrlees (1971), and the more recent literature that spun from the work of Saez (2001). Similarly, we are certainly not the first to incorporate the costs of redistribution into the analysis of the welfare effects of policies. The need to do so was actually anticipated by Hicks in the concluding passages of his 1939 paper, and was subsequently explored by Kaplow (2004) and, more recently, by Hendren (2014). Hendren (2014), in particular, estimates the *inequality deflator* associated with the transfer of one dollar of income from individuals at different positions in the U.S. income distribution to the rest of the U.S. population. He finds that this deflator is higher for rich individuals than for poor individuals and uses it to quantify the effects of increased income inequality on U.S. economic growth. His approach to costly redistribution is certainly more sophisticated than the one adopted in this paper, as it involves an estimation of the joint distribution of marginal tax rates and the income distribution using the *universe* of U.S. income tax returns in 2012. The thought experiment that motivates his work is however distinct from ours. While we seek to understand the efficiency costs associated with the behavioral responses of agents triggered by trade-induced shifts across marginal tax rates, his focus is on understanding the efficiency consequences of *local changes* to the nonlinear income tax schedule aimed at compensating the losers from a particular policy. It might be fruitful to adopt his approach to the study of the effects of trade liberalization, but we leave this for future research.

2 Inequality and Welfare: A Primer

We begin our analysis in this section by considering various approaches to measuring the evolution of social welfare in the face of changing inequality and when complete and costless redis-

[9] Davidson and Matusz (2006) design the lowest cost compensation policies for the losers from trade in a two-sector economy with heterogenous agents and participation decisions, but fixed labor supply (see also Egger and Kreickemeier, 2009b).

[10] Rodrik (1992) is a noteworthy antecedent to our work in discussing the costs of redistribution following changes in trade policy.

tribution is infeasible. We first review the Kaldor-Hicks principle and the Atkinson's welfarist approach, and then present our costly redistribution approach. While doing so, we introduce our two main inequality correction terms for measuring welfare gains—the welfarist correction and the costly-redistribution correction—and discuss their properties. In order to simplify the exposition, the framework developed in this section will leave some of the primitive determinants of income, welfare and costly redistribution unspecified. In section 3, we formalize these correction terms in a context of a simple yet fully microfounded general-equilibrium model, and we illustrate how to use this framework to provide back-of-the-envelope calculations of welfare changes when both aggregate income and inequality change over time.

2.1 Economic Environment

Consider a society composed of a measure one of individuals indexed by their ability level φ with associated real earnings r_φ. Agents' preferences are represented by a utility function u defined over consumption c_φ. Agents' consumption is in turn equal to their real *disposable* income r_φ^d, defined as:

$$r_\varphi^d = \left[1 - \tau(r_\varphi)\right] r_\varphi + T_\varphi, \tag{1}$$

where $\tau\left(r_\varphi\right)$ denotes a non-linear income tax and T_φ represents a lump-sum transfer. The distribution of φ in the population is given by the cumulative distribution function H_φ, while the associated income distribution for real before-tax earnings is denoted by F_r. For simplicity, we assume –for the time being– that the government budget is balanced so that

$$\int r_\varphi^d \mathrm{d}H_\varphi = \int r \mathrm{d}F_r = R.$$

The society is evaluating the consequences of a policy (such as a trade liberalization) that would generate heterogeneous effects on agents' real incomes, thereby leading to a shift from the initial distribution of earnings F_r to a new distribution of real income F_r'. What are the welfare consequences of the move from F_r to F_r'?

We discuss below three different approaches to the evaluation of the social welfare implications of the move from F_r to F_r'.

2.2 The Kaldor-Hicks Principle

The Kaldor-Hicks compensation principle constitutes the standard approach to evaluating the welfare effects of a policy. To identify a Kaldor-Hicks improvement, one starts by computing the compensating or equivalent variation for each individual associated with the particular policy under study, and these money metrics are then aggregated across all individuals. In our example above, this principle implies that *mean* real income growth is a sufficient statistic for comparing social welfare under F_r and F_r', regardless of the effect of the policy on the higher moments of the income distribution.

Let us illustrate this for the case of the compensating variation, which we denote with v_φ

for an individual of type φ and is defined by:

$$u\left(r_{\varphi}^{d\prime}+v_{\varphi}\right)=u\left(r_{\varphi}^{d}\right). \tag{2}$$

It follows that the required aggregate compensation satisfies:

$$-\int v_{\varphi}\mathrm{d}H_{\varphi}=\int r_{\varphi}^{d\prime}\mathrm{d}H_{\varphi}-\int r_{\varphi}^{d}\mathrm{d}H_{\varphi}=\int r\mathrm{d}F_{r}^{\prime}-\int r\mathrm{d}F_{r}. \tag{3}$$

Clearly, the right-hand-side of (3) corresponds to the change in aggregate real income, which we write as $R'-R$. If this quantity is positive, it means that the amount of money necessary to restore the losers' welfare to its pre-policy level is lower than the amount that winners are jointly willing to give up for the policy to be adopted. In order to quantify the gains from trade, it is standard to express the change in (3) as a percentage change relative to the initial level of aggregate real income R, which we can denote by

$$\left.\frac{W'}{W}\right|_{\text{Kaldor-Hicks}}=1+\mu^{R}\equiv\frac{R'}{R}. \tag{4}$$

The welfare gains from the policy thus correspond to the mean real income growth μ^R it generates. More generally, the overall welfare impact of other exogenous shocks can be evaluated analogously by only considering their effect on average income (or GDP).

Although we have assumed that all agents have a common indirect utility function U, it is clear from equation (2) that the result in (3) will apply even when agents are heterogeneous not only in income but also in preferences. This is a key appealing feature of the Kaldor-Hicks criterion: it does not rely on interpersonal comparisons of utility.[11]

As noted in the Introduction, there are however two key limitations of the Kaldor-Hicks criterion. First, the fact that there is the *potential* for the winners to compensate the losers does not mean that this compensation will *actually* takes place in practice. If little redistribution takes place and the ex-post distribution of income is much more unequal than the ex-ante one, it is less clear that mean income should be a sufficient statistic for measuring welfare changes. Second, the focus on compensating or equivalent variations is justified only in the presence of lump-sum taxes, which ensure a frictionless redistribution of gains across the individuals. While a useful theoretical tool, lump-sum transfers are informationally intensive and rarely feasible in practice. Naturally, compensation may also be achievable via other forms of redistribution, but these alternative instruments are likely to impact economic efficiency and thus the magnitude of the welfare gains from a policy.

In light of these limitations, we next discuss two alternative (and complementary) approaches to policy evaluation that explicitly correct for the induced effect of a policy on income inequality.

[11] In other words, the welfare gains in (4) are independent of the particular cardinal utility functions that are chosen to represent the ordinal preferences of individuals.

2.3 The Welfarist Approach

The welfarist (or social welfare) approach to policy evaluation begins by positing the existence of a social welfare function that maps the vector of agents' welfare levels into a single real number. It is customary to express this function as an integral of concave transformations of agents' *actual* (and not potential) disposable incomes (and thus consumption levels):[12]

$$V = \int u(r_\varphi^d)\mathrm{d}H_\varphi, \tag{5}$$

where $u'(\cdot) > 0$ and $u''(\cdot) \leq 0$. There are at least two possible justifications for specifying $u(\cdot)$ as a concave function. First, given two distributions of *disposable* income with the same mean, one would expect society to prefer the one with the lowest dispersion or inequality (cf. Atkinson, 1970), with the concavity of $u(\cdot)$ reflecting inequality aversion on the part of the social planner. It is important to emphasize that, under the plausible assumption that agents' preferences feature diminishing marginal utility of income, inequality aversion is completely consistent with a utilitarian social planner that simply seeks to maximize the sum of agent's utilities. A second justification for the concavity of $u(\cdot)$ is that it might capture risk aversion on the part of ex-ante identical individuals in some sort of "original position" attempting to compute the individual welfare implications of changes in the environment behind a "veil of ignorance" (cf., Vickrey, 1945; Harsanyi, 1953).[13]

To fix ideas, we shall follow Atkinson (1970) and consider a constant-elasticity function:

$$u(r_\varphi^d) = \frac{\left(r_\varphi^d\right)^{1-\rho} - 1}{1-\rho}, \tag{6}$$

where $\rho \geq 0$ can be interpreted as reflecting a constant degree of inequality aversion on the part of the social planner or a constant degree of risk aversion on the part of agents in the original position (or a combination of both). In order to express social welfare changes in terms of aggregate consumption equivalent changes, it will further prove convenient to consider the simple monotonic transformation

$$W = \left[1 + (1-\rho)V\right]^{1/(1-\rho)} \tag{7}$$

of the social welfare function in (5). With this transformation, social welfare can be expressed as a multiplicatively separable function of aggregate real income R and a term Δ, which is inversely related to the level of inequality underlying the distribution of disposable income:

$$W = \Delta \times R, \tag{8}$$

[12] More generally, the social welfare function is represented by a vector (density function) of utility weights which the planner uses to aggregate individual welfare levels.

[13] This assumes that agents are not able to ex-ante insure against this ex-post dispersion in income. We view trade shocks as particularly hard to insure against. Furthermore, as mentioned in the Introduction, at least in the U.S., consumption inequality appears to have tracked income inequality in recent decades.

where

$$\Delta = \Delta(F_r^d, \rho) = \frac{\left[\mathbb{E}\left(r_\varphi^d\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}}{\mathbb{E}r_\varphi^d}. \tag{9}$$

The term Δ, which we will refer to as a *welfarist inequality correction*, corresponds *exactly* to one minus the Atkinson (1970) index, a widely used measure of inequality. By Jensen's inequality we have that $\Delta \leq 1$, with $\Delta = 1$ only if either there is no inequality aversion ($\rho = 0$) or if the distribution of disposable income F_r^d is fully egalitarian (has zero dispersion). Furthermore, Δ tends to be lower, the higher is the level of inequality in the distribution of income F_r^d or the higher is inequality aversion ρ. To be more precise, while Δ is invariant to proportional changes of the income distribution (i.e., when all income levels are scaled by the same constant), Δ is reduced by mean-preserving spreads of the distribution of disposable income (cf., Atkinson, 1970). And holding constant the distribution of disposable income, F_r^d, the higher is the degree of inequality (or risk) aversion ρ, the greater is the correction and the smaller is Δ (see Appendix A.1 for a formal proof). As we show in Appendix A.2, for certain often-used distributions of income, it is also possible to relate Δ to the Gini coefficient associated with F_r^d.

The expression for welfare (8) immediately implies that the percentage welfare gains from a policy are given by:

$$\left.\frac{W'}{W}\right|_{\text{Welfarist}} = (1 + \mu^R) \times \frac{\Delta'}{\Delta}, \tag{10}$$

where μ^R is the growth rate of real income as defined in (4) and $\Delta' = \Delta(F_r^{d\prime}, \rho)$ corresponds to the correction term under the new income distribution. Thus in the absence of any effect of trade on inequality as captured by Δ, the change in welfare corresponds exactly to the percentage change in real income μ^R, as in the Kaldor-Hicks compensation principle approach in (4). Nevertheless, if trade increases inequality, then welfare increases by less than $1 + \mu^R$, with a larger downward correction the larger is ρ and, of course, the larger the increase in inequality. The particularly size of the correction can be easily computed with data (real or counterfactual) on the distribution of disposable income before and after the policy, as we shall illustrate in sections 4 and 6.

As mentioned above, an advantage of using the function W in (7) instead of any other monotonic transformations of V in (5) is that the change in welfare in (10) also corresponds to the consumption-equivalent change in social welfare of moving from F_r^d to $F_r^{d\prime}$. More specifically, it is easy to verify (see Appendix A.1) that if one were to compute the percentage change in all agents' consumption or disposable income that would make society indifferent between F_r^d and $F_r^{d\prime}$, the answer one would get would be $\mu^C = (1 + \mu^R) \times \frac{\Delta'}{\Delta} - 1$ regardless of whether social welfare is measured in terms of the function V or of any monotonic transformation of V (including W).

2.4 The Costly-Redistribution Approach

Despite its widespread use in the optimal policy literature, the welfarist approach remains controversial. This is in large part due to the sensitivity of its prescriptions to the value of certain parameters, such as the degree of inequality aversion (or, more generally, the social marginal weights assigned to agents with different income), that are difficult to measure and over which people might have vastly different ethical views.

We next consider a complementary approach that is more akin to the Kaldor-Hicks compensation principle, but that explicitly models the fact that redistribution is costly, with the costs of redistribution are increasing in the extent of economic inequality. The welfare correction in this case quantifies the forgone gains in real income due to the costly redistribution mechanism put in place by society to reduce income inequality.

For this purpose, we return to our previous example but now assume that lump-sum transfers are not feasible (i.e., $T_\varphi \equiv 0$) and redistribution has to work through the income tax system. Above, we have introduced a general nonlinear income tax $\tau(r_\varphi)$, but we will now focus on the particular case, used among others by Benabou (2002) and Heathcote, Storesletten, and Violante (2014), in which

$$r_\varphi^d = \left[1 - \tau(r_\varphi)\right] r_\varphi = k r_\varphi^{1-\phi}, \tag{11}$$

for some constant k that can be set to ensure that the government budget is balanced. Average net-of-tax rates thus decrease in reported income at a constant rate ϕ, with this parameter governing the degree of progressivity of the tax system. When $\phi = 0$, all agents face the same tax rate k and there is no redistribution from the rich to the poor; in fact, with budget balance there is no redistribution whatsoever. When $\phi = 1$, (11) implies that all agents end up with the same after-tax income and thus redistribution is full and eliminates inequality.[14]

The specification in equation (11) may seem quite ad hoc and unlikely to provide a valid approximation to the complicated tax and transfer systems employed in modern economies. Nevertheless, its implied log-linear relationship between market income and income after taxes and transfers fits U.S. data remarkably well, as we will illustrate in more detail in section 4 (see also Heathcote, Storesletten, and Violante, 2014 and Guner, Kaygusuz, and Ventura, 2014).

A larger degree of progressivity tends to compress the after-tax income distribution, but it implies that rich people face disproportionately larger marginal tax rates. More specifically, the marginal tax rate implied by (11) is given by $\tau^m(r_\varphi) = 1 - k\,(1-\phi)\, r_\varphi^{-\phi}$ and thus rises with both the degree of tax progressivity ϕ as well as the level of income r_φ. To the extent that higher marginal tax rates generate behavioral responses of agents that lead them to generate less income than they would under a lower marginal tax rate, the increased redistribution brought about by a higher degree of progressivity will generate costs. To capture this costly aspect of redistribution in a simple though fairly standard way, we posit the existence of a a positive, constant elasticity of taxable (realized) income to the net-of-marginal-tax rate $1 - \tau^m$:

$$\varepsilon \equiv \frac{\partial \log r_\varphi}{\partial \log(1 - \tau^m(r_\varphi))} \geq 0, \tag{12}$$

[14] More generally, the results in Jakobsson (1976) and Kakwani (1977) imply that, starting from a fixed arbitrary distribution of pre-tax income, an increase in ϕ necessarily leads to a more egalitarian distribution of after-tax income, in the sense that it makes disposable income more evenly distributed according to the Lorenz criterion.

where $\tau^m(r_\varphi)$ is the marginal tax rate faced by agents with income r_φ.

The combination of a progressive tax system of the type in (11) and a positive elasticity of taxable income ε makes redistribution from rich people to poor people costly, thereby motivating an alternative correction to the standard measures of the welfare effects of a policy. More specifically, one can manipulate equations (11) and (12) and impose budget balance, to obtain:

$$R = \Theta \times \tilde{R}, \tag{13}$$

where $\tilde{R}$ is the *potential* income in the absence of progressive redistribution (i.e., counterfactual income obtain when setting $\phi = 0$ in the tax schedule), and

$$\Theta = \Theta(F_r, \phi, \varepsilon) = (1-\phi)^\varepsilon \frac{\left(\mathbb{E}r_\varphi\right)^{1+\varepsilon}}{\left(\mathbb{E}r_\varphi^{1-\phi}\right)^\varepsilon \cdot \left(\mathbb{E}r_\varphi^{1+\varepsilon\phi}\right)} \tag{14}$$

is a term we refer to as our *costly-redistribution inequality correction.*

Although perhaps not immediate from inspection of (14), Hölder's inequality implies that the second term is no larger than 1, which in turn implies $\Theta \leq 1$. Furthermore, $\Theta = 1$ if and only if the tax-transfer system features zero progressivity ($\phi = 0$) or if the elasticity of taxable income is zero ($\varepsilon = 0$). It thus follows that, when $\phi > 0$ and $\varepsilon > 0$, real income is lower than it would be in the absence of distortionary redistribution. In fact, in Appendix A.1 we show that Θ is strictly decreasing in the tax progressivity rate ϕ for any primitive distribution of potential output $\tilde{r}_\varphi$, thus formalizing the efficiency costs of enhancing redistribution. Holding the other parameters constant, the term Θ also depends on the *primitive* degree of income inequality: Θ is highest whenever the income distribution is perfectly egalitarian and it tends to be lower the more unequal is the distribution of income. More specifically, when considering two distributions of income F_r and F_r', it is easy to show that $\Theta(F_r', \phi, \varepsilon) < \Theta(F_r, \phi, \varepsilon)$ when F_r' is a mean preserving multiplicative spread of F_r.[15] Conversely, Θ is invariant to proportional changes of the income distribution (i.e., when all income levels increase proportionately). In analogy to the Atkinson index, one can interpret Θ as a complementary welfare-relevant measure of inequality, and for certain standard distributions, Θ can be related to the Gini coefficient associated with F_r (see Appendix A.2 for details).

We are now ready to revisit our initial question of how should society evaluate the welfare implications of the move from F_r to F_r'. Even when one adheres to a welfare criterion, such as the Kaldor-Hicks principle, that judges policies based on their implications for real income growth, with costly redistribution society will take into account the effects of the policy on higher moments of the income distribution. The reason for this is that, in the absence of lump-sum transfers, those higher moments shape the determination of mean disposable income. More

[15] The distribution F_r' is a mean preserving multiplicative spread of F_r whenever there exists a random variable θ independent of the original income r such that $r' = (1+\theta)\, r$ with $\mathbb{E}(\theta) = 0$. Note that Θ is less than one even when all agents share the same income and thus there is no redistribution in equilibrium. The reason for this is that when considering an off-the-equilibrium path deviation that would increase an agent's income, this agent understands that it will be taxed as a result of that deviation. This is captured by the term $(1-\phi)^\varepsilon$ in (14).

precisely, building on (13) we can express the aggregate real income gains of the policy as

$$1+\mu^R = \left.\frac{R'}{R}\right|_{\text{Costly Red.}} = (1+\tilde{\mu}^R)\times\frac{\Theta'}{\Theta}, \tag{15}$$

where $1+\tilde{\mu}^R \equiv \tilde{R}'/\tilde{R}$ measures the real income gains in the absence of costly redistribution. Whenever the policy has no measurable impact on Θ, the change in welfare corresponds exactly to real income growth of a hypothetical *Kaldor-Hicks economy* that could use lump-sum transfers for redistribution purposes. Such an equivalence would hold when the policy increases the incomes of all agents proportionately (and ϕ and ε do not change). If however the policy increases inequality and thereby lowers Θ, the implied change in aggregate income will be strictly lower than in the case in which inequality had remained unaffected. To summarize, the costly redistribution correction measures the forgone gains in real income due to the interaction between the increased inequality and the progressivity of the tax schedule.

Although we hope that the discussion in this section has served a useful pedagogical role, a proper analysis of how the welfarist and costly redistribution corrections shape social welfare requires the development of a fully specified model in which the income distribution is endogenized and in which the response of agents to taxation is microfounded and taken into account in computing social welfare. We turn to this task in the next section.

3 Inequality and Welfare in a Constant-Elasticity Model

In this section, we develop a simple general equilibrium framework, which specifies how the ability of individuals and their labor supply decisions translate in equilibrium earnings and welfare levels given the tax system in place. In light of our choices of functional forms, we refer to our model as the constant-elasticity model.

The model features four constant elasticity parameters, which we introduce below: (i) a constant Frisch elasticity of labor supply ($1/(\gamma-1)$); (ii) a constant elasticity of substitution between the labor services (or tasks) performed by different agents in society ($1/(1-\beta)$); (iii) a constant degree of tax progressivity (ϕ); and (iv) a constant social inequality aversion (ρ). This constant-elasticity structure results in a tractable general equilibrium characterization, which is particularly useful to illustrate our welfare corrections terms. We should emphasize, however, that our model will place little structure on the underlying primitive distribution of ability, and can thus flexibly accommodate *any* equilibrium distribution of income one may choose to calibrate the model to. Let us next introduce the key ingredients of the model more formally.

3.1 Preferences, Technology and Individual Behavior

Consider for now a closed economy inhabited by a continuum of agents with GHH preferences (cf., Greenwood, Hercowitz, and Huffman, 1988) over the consumption of an aggregate good c

and labor ℓ:

$$u(c,\ell) = c - \frac{1}{\gamma}\ell^{\gamma}. \tag{16}$$

The parameter $\gamma \geq 1$ controls the Frisch elasticity of labor supply, which is given by $1/(\gamma-1)$ and is decreasing in γ. In the presence of elastically supplied labor, theoretically-grounded measures of welfare need to correct income for the disutility costs of producing it, an issue we ignored in section 2. This utility specification results in no income effects on labor supply and is often adopted in the optimal taxation literature.

Each individual produces output $y = \varphi\ell$ of his own variety of a task (or intermediate good) where φ is individual ability and is distributed according to H_φ as in section 2. The tasks performed by different agents are imperfect substitutes and are combined in the production of the aggregate consumption (final) good according to

$$Q = \left(\int y_\varphi^\beta \mathrm{d}H_\varphi\right)^{1/\beta},$$

where $\beta \in (0,1]$ is a parameter that controls the elasticity of substitution $1/(1-\beta)$ across tasks. In the limiting case of $\beta = 1$, the individual tasks become perfect substitutes, and the model turns into a special case of a neoclassical Mirrlees (1971) economy. Imperfect substitutability becomes essential when we introduce an explicit model of international trade in section 5, but for the qualitative implications of this section whether $\beta = 1$ or $\beta < 1$ is not important.

Under the above assumptions, the market (real) earnings of an individuals supplying y units of his task to the market are given by:[16]

$$r = Q^{1-\beta}y^\beta. \tag{17}$$

Notice that when $\beta < 1$, the demand for each individual task is increasing in aggregate income

$$Q = R = \int r_\varphi \mathrm{d}H_\varphi,$$

yet the agents face decreasing demand schedules and as a result their revenues are concave in their own output. When $\beta = 1$, the individual revenues are simply $r = y$, and thus are only a function of their ability and labor supply decisions.

Individual consumption equals after-tax income, $c = r^d = [1-\tau(r)]r$. As in section 2, we assume that the tax-transfer system is well approximated by equation (11), where the parameter ϕ governs tax progressivity and the parameter k controls the average tax rate across agents. The government uses collected taxes for redistribution and to finance exogenous government spending G, and runs a balanced budget. In other words, the total income of the economy equals the

[16]The demand for an individual task variety is given by $q = Q(p/P)^{-\frac{1}{1-\beta}}$, were p is the price of the variety and $P = \left(\int p_\varphi^{-\frac{\beta}{1-\beta}} \mathrm{d}H_\varphi\right)^{(1-\beta)/\beta}$ is the price of the final good. We normalize $P = 1$ so that all nominal quantities in the economy are in terms of the final good, and thus are in real terms as well. Under these circumstances, task revenues are $r = pq = Q^{1-\beta}y^\beta$, where we have substituted the market clearing condition $q = y$.

sum of total private consumption (aggregate disposable income) and government spending, so $Q = \int r_\varphi^d \mathrm{d}H_\varphi + G$. We further assume that government spending is a fraction g of GDP, i.e. $G = gQ$, and it does not directly affect the individual utilities in (16). Under these circumstances, we can rewrite the government budget balance as:

$$k \int r_\varphi^{1-\phi} \mathrm{d}H_\varphi = (1-g)Q, \tag{18}$$

which defines a relationship between k and g given the tax schedule progressivity ϕ. In other words, given the exogenous share of government spending g, there exists a unique average tax parameter k which balances the government budget for any given level of tax progressivity ϕ.

Individuals maximize utility (16) by choosing their labor supply and consuming the resulting disposable income, a program that combining (11) and (17) we can write as:

$$u_\varphi = \max_\ell \left\{ k \Big[Q^{1-\beta} (\varphi \ell)^\beta \Big]^{1-\phi} - \frac{1}{\gamma} \ell^\gamma \right\}.$$

The solution for equilibrium revenues and utilities is given by:

$$r_\varphi = \Big[\beta(1-\phi)k \Big]^{\frac{\varepsilon}{1+\varepsilon\phi}} \Big[Q^{1-\beta} \varphi^\beta \Big]^{\frac{1+\varepsilon}{1+\varepsilon\phi}}, \tag{19}$$

$$u_\varphi = \frac{1+\varepsilon\phi}{1+\varepsilon} k r_\varphi^{1-\phi}, \tag{20}$$

where we have made use of the following auxiliary constant:

$$\varepsilon \equiv \frac{\beta}{\gamma - \beta},$$

which also equals the overall elasticity of taxable income to changes in marginal tax rates, as previously defined in (12).[17] When tasks are perfectly substitutable ($\beta = 1$), this elasticity ε coincides with the Frisch elasticity of labor supply $1/(\gamma-1)$. Yet with imperfect substitutability in tasks ($\beta < 1$), this elasticity is reduced by the downward pressure of increased output on prices.

Equations (19)–(20) show how individual ability translates into equilibrium market revenue and individual utility. The latter is proportional to after-tax income because the utility cost of labor effort is proportional to disposable income under the optimal allocation. Equilibrium revenues are a power transformation of underlying individual abilities, with the power increasing in the elasticity parameters ε and β, and decreasing in the progressivity of taxation ϕ. Tax progressivity not only reduces the dispersion of after-tax incomes and utilities, but also compresses the distribution of pre-tax market revenues as it has a disincentive effect on labor supply, which is particularly acute for high-ability individuals facing higher marginal tax rates.

[17] To see this, remember that the marginal tax rate associated with (11) is given by $\tau^m(r_\varphi) = 1 - k(1-\phi) r_\varphi^{-\phi}$. Plugging this marginal tax rate into (19) and simplifying delivers $r_\varphi = (\beta(1-\tau^m(r_\varphi)))^\varepsilon \left(Q^{1-\beta}\varphi^\beta\right)^{1+\varepsilon}$.

3.2 Aggregate Income and Social Welfare

The characterization of equilibrium revenues and utilities relies on two endogenous aggregate variables, k and Q. The closed-form solutions for these variables are provided in Appendix A.3, where we show that aggregate income (GDP) of the economy can be expressed as

$$Q = \Theta^{\kappa} \tilde{Q}, \tag{21}$$

where $\Theta < 1$ is the same costly-redistribution correction term introduced above in equation (14), and where

$$\tilde{Q} = [\beta(1-g)]^{\kappa\varepsilon} \left(\int \varphi^{\beta(1+\varepsilon)} \mathrm{d}H_{\varphi} \right)^{\kappa}$$

is the counterfactual (potential) aggregate real GDP with a flat tax schedule characterized by $\phi = 0$ and $k = 1 - g$ to finance government spending. In these expressions, the auxiliary parameter κ

$$\kappa \equiv \frac{1}{1 - (1-\beta)(1+\varepsilon)} \geq 1$$

captures an amplification effect associated with the aggregate demand externality (or love-for-variety effect) stemming from the imperfect substitutability of tasks.[18]

Several comments are in order. First, note that aggregate real income in equation (21) depends on the costly redistribution correction term Θ and on potential real income, which is a simple function of the primitive fundamentals of the model, namely the ratio g of government spending to GDP, the distribution of ability H_{φ}, the task-substitutability parameter β, and the Frisch elasticity of labor supply γ (which together with β determine the elasticity of taxable income ε). Second, in the absence of progressive taxation, realized and potential GDP coincide because remember that when $\phi = 0$, $\Theta = 1$. Third, note that (21) is the counterpart to equation (13) in section 2, with the only difference being that now the output loss $Q/\tilde{Q}$ is amplified by the aggregate demand externality, which manifests itself in the exponent $\kappa > 1$ on Θ, and operates in the model whenever $\beta < 1$.

So far, we have focused on a discussion of the determination of aggregate real income in the model. Equation (20) provides the utility level associated with the disposable income and labor supply decisions of an individual with ability φ. In order to aggregate these utility levels into a measure of social welfare, we adopt the welfarist approach and express social welfare as

$$W = \left(\int u_{\varphi}^{1-\rho} \mathrm{d}H_{\varphi} \right)^{\frac{1}{1-\rho}}, \tag{22}$$

which is the exact counterpart to our earlier equation (7). Note that the risk aversion parameter $\rho \geq 0$ is inconsequential for the choices of individuals in this static model, and only matters for cross-individual welfare comparisons. Therefore, ρ can be viewed as either the property of

[18] Note that when $\beta = 1$, $\kappa = 1$, but κ is otherwise increasing in ε and decreasing in β. As is clear from the definition of κ, we need to impose the stability condition $(1-\beta)(1+\varepsilon) < 1$, which is satisfied if ε is not too large or β is not too small.

individual utilities of the agents or the social inequality aversion parameter. When $\rho = 0$, social welfare corresponds to the simple integral of utility levels across individuals, which remember are linear in real disposable income.

This completes the description of the model environment, and we can now characterize equilibrium welfare W given the solution for equilibrium utilities in (20). We do this in two steps. First, we characterize (see Appendix A.4 for a proof):

Proposition 1. *The welfare in the economy with zero tax progressivity ($\phi = 0$) and no inequality aversion ($\rho = 0$) is given by:*

$$\tilde{W} = \frac{1-g}{1+\varepsilon} \times \tilde{Q}. \tag{23}$$

Note that welfare in this case is closely related to real GDP, with the (welfare) cost of producing GDP captured by the discount term in front of $\tilde{Q}$ in (23). The numerator $1-g$ of this term reflects the share of the output of the economy that goes towards the provision of the public good, which is financed via a proportional tax schedule $k = 1-g$ when $\phi = 0$.[19] The denominator $1+\varepsilon$ reflects the disutility costs of producing the output $\tilde{Q}$. The immediate corollary of Proposition 1 is that in the absence of inequality aversion and tax progressivity, changes in welfare can be measured using the growth rate of GDP

$$\frac{\tilde{W}' - \tilde{W}}{\tilde{W}} = \tilde{\mu}^R = \frac{\tilde{Q}' - \tilde{Q}}{\tilde{Q}},$$

provided that the elasticity ε and the share of public spending g stay constant over time.

This result illustrates that a criterion analogous to the Kaldor-Hicks prescription in (4) may still apply in more general settings, even when lump-sum taxes are unavailable and average taxes are positive, provided that society does not care about inequality and does not use a progressive tax system to address it.

Nevertheless, outside this limiting case with $\phi = \rho = 0$, real income growth is no longer an appropriate measure of welfare gains, and instead we have (see Appendix A.4 for a proof):

Proposition 2. *Outside the case $\phi = \rho = 0$, social welfare can be written as:*

$$W = \Delta \times (1+\varepsilon\phi)\Theta^{\kappa} \times \tilde{W}, \tag{24}$$

where Δ and Θ are the welfarist and the costly-redistribution corrections defined in (9) *and* (14)*, respectively.*

Note that the two inequality correction terms that were introduced earlier in section 2 appear explicitly in the welfare expression in (24). Indeed, realized welfare W equals potential welfare $\tilde{W}$ discounted in turn by the two correction terms, which are both less than 1 (see our discussion in section 2).

[19] Remember that potential output is given by $\tilde{Q} = [\beta(1-g)]^{\kappa\varepsilon}\left(\int \varphi^{\beta(1+\varepsilon)} \mathrm{d}H_{\varphi}\right)^{\kappa}$, and hence is itself decreasing in $k = 1-g$ due to the disincentive effect of the average tax rate g even in the absence of progressivity of taxation.

The effect of inequality aversion on social welfare is captured by the exact same term Δ derived in section 2, and is closely related to the Atkinson inequality measure. The effect of costly redistribution is instead slightly modified relative to our previous derivations in section 2. First, and as already discussed above, the presence of aggregate demand externalities magnify the loss of output associated with distortionary taxation, and $\Theta < 1$ is now raised to a power $\kappa > 1$. Second, note that (24) incorporates a new term $(1 + \varepsilon\phi)$ which captures the fact that lower output comes along with a lower disutility of labor effort, which other things equal, raises welfare. Despite the presence of this term, the overall effect of distortionary taxation captured by the term $(1 + \varepsilon\phi)\Theta^{\kappa}$ inherits the same properties of the term Θ described in section 2. In particular, this term can be split into the product of $(1 + \varepsilon\phi)(1 - \phi)^{\kappa\varepsilon}$ and $\left(\mathbb{E}r_{\varphi}\right)^{\kappa(1+\varepsilon)} / \left[\left(\mathbb{E}r_{\varphi}^{1-\phi}\right)^{\varepsilon} \cdot \left(\mathbb{E}r_{\varphi}^{1+\varepsilon\phi}\right)\right]$, with both of these terms being strictly less than 1 when $\phi > 0$ and the dispersion of income is positive. The first term captures the utility loss from taxation in the absence of inequality (and continues to be decreasing in ε and ϕ), while the second term captures the additional loss due to the interaction of inequality with a progressive income tax schedule.

Although this is not the focus of this paper, the welfare decomposition in equation (24) captures the tradeoff faced by a benevolent government (maximizing social welfare W) when deciding on the degree of progressivity ϕ of the tax system. Because the welfarist correction Δ is negatively affected by an increase in inequality in disposable income, this correction term is increasing in ϕ. Conversely the costly-redistribution correction (and also the full correction term $(1 + \varepsilon\phi)\Theta^{\kappa}$) is decreasing in ϕ, due to the higher marginal tax rates associated with a more progressive tax system. Therefore, when setting the optimal ϕ, the government necessarily balances these two conflicting forces (see Appendix A.1 for details).

To summarize, we have shown that social welfare can be expressed as a multiplicatively separable function of three terms: (i) potential welfare in a hypothetical (Kaldor-Hicks) world with non-distortionary taxation, discounted by (ii) our welfarist correction, and (iii) our (modified) costly-redistribution correction. The presence of these two correction terms introduces a tradeoff for the policy maker when deciding on the optimal degree of tax progressivity. Independently of the amount of redistribution that society chooses to implement, these two corrections reduce welfare disproportionately more in environments with higher economic inequality. In the model developed so far such increases in inequality can only originate from increases in the dispersion of ability across agents (perhaps due to skill-biased technological change) or from increases in the primitive parameters β and γ. In section 5, we will show, however, that trade integration can generate qualitatively similar effects. Before doing so, and to build some intuition for our quantitative analysis, in the next subsection we provide a preliminary look at U.S. data through the prism of our closed-economy model.

4 A Preliminary Look at the Data

Although the main goal of this paper is to apply the tools developed so far to the study of the welfare gains from trade integration, in this section we take a brief detour to illustrate the usefulness of our closed-economy model in interpreting the consequences of the observed rise in inequality in the U.S. in recent times.

More specifically, in this section we decompose social welfare in the U.S. over the period 1979-2007 according to equation (24), thus backing out the size and evolution of the welfarist and inequality correction terms. We then use this expression to compute the income and welfare levels that would have attained in counterfactual scenarios in which U.S. income inequality had not increased as much as it did over this period.

4.1 Calibration

In order to put the above model to work, it is necessary to calibrate its key parameters. Remember that the primitive parameters of the model are the Frish elasticity parameter γ, the task substitutability parameter β, the degree of tax progressivity ϕ, the share of government spending in GDP g, the distribution of ability in society H_{φ}, and the inequality aversion parameter ρ. Some of these objects, such as the distribution of agents' ability, are notoriously difficult to calibrate. Fortunately, we shall see that, for our purposes, it will suffice to calibrate (i) the degree of tax progressivity of income ϕ, (ii) the distribution of market income r_{φ}, (iii) the elasticity of taxable income $\varepsilon = \beta/(\gamma - \beta)$, and (iv) the degree of substitutability between the tasks provided by different workers, as captured by β. Let us discuss each of these in turn.

Consider first our modeling of the tax-transfer system in equation (11). This specification may seem quite ad hoc, but the log-linear relationship between market income and income after taxes and transfers implied by equation (11) fits U.S. data remarkably well. This is illustrated in Figure 2 using CBO data for eight percentiles of the income distribution for various years in the period 1979-2007.[20] The best log-linear fit of the data achieves a remarkable R-squared of 0.983 or higher in all years.

Inspection of the different panels of Figure 2 suggests that the degree of tax progressivity appears to be lower in recent years than at the beginning of the period. This is more clearly illustrated in Figure 3, which reports the estimate of ϕ year by year. Given the remarkable fit of equation (11), we will use these yearly estimates of ϕ to calibrate the time path of progressivity over the period 1979-2007. Heathcote, Storesletten, and Violante (2014) find an equally good fit of this log-linear relationship with data from the Panel Study of Income Dynamics (PSID) for survey years 2000, 2002, 2004, and 2006, with an implied value $\phi = 0.151$, which is very much consistent with the estimates we obtain from CBO during the period 2000-06.

We next turn to our calibration of the distribution of market income. Because we are wary that our quantitative results might be sensitive to fine features of the income distribution (such

[20] In particular, each panel of Figure 2 depicts market income and income after taxes and transfers for the first four quintiles of the income distribution, as well as the 81st to 90th percentiles, the 91st to 95th percentiles, the 96th to 99th percentiles, and the top 1 percent, for the period 1970-2007. Market income consists of labor income, business income, capital gains (profits realized from the sale of assets), capital income (excluding capital gains), income received in retirement for past services, and other sources of income. Government transfers include cash payments and in-kind benefits from social insurance and other government assistance programs. Federal tax liabilities include individual income taxes, social insurance or payroll taxes, excise taxes, and corporate income taxes.

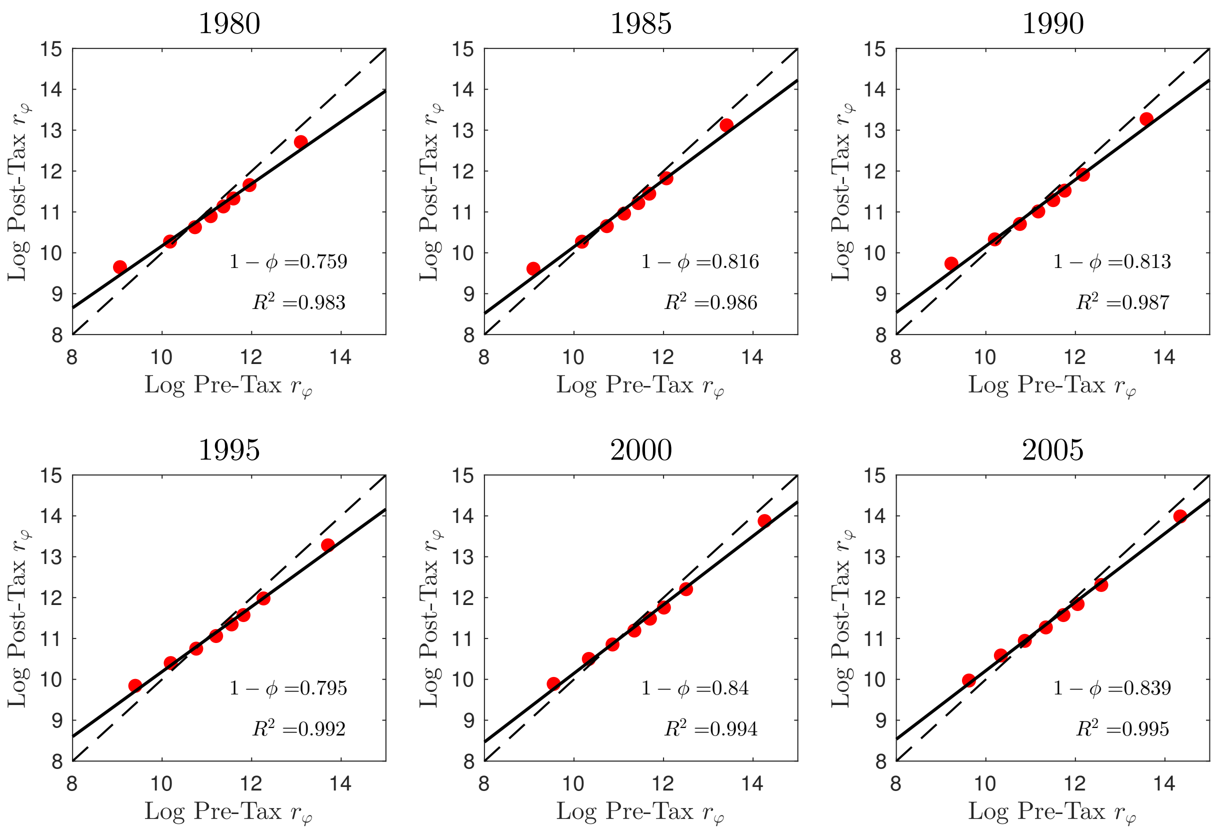

Figure 2: Progressivity

as the shape of its right tail), we deem it necessary to seek richer information on the U.S. income distribution than that provided by the CBO data we used to calibrate ϕ. Following recent empirical work on top income levels (e.g., Piketty and Saez, 2003), we employ the public use samples of U.S. Federal Individual Income Tax returns available from the NBER website. These amount to approximately 3.5 million anonymized tax returns (about 150,000 per year) over the period 1979-2007. Contrary to survey-based sources of income distribution data, the NBER IRS data is more likely to provide an accurate picture of the income of particularly rich taxpayers. To ensure the representativeness of the sample, we further apply the sampling weights provided by the NBER. We map before-tax income r_φ in the model to adjusted gross income (AGI) in line 37 of IRS Form 1040 and restrict the sample to returns with a strictly positive AGI. Together with our yearly estimates of ϕ, it is then possible to estimate disposable income r_φ^d up to a constant (k) which is irrelevant for the computation of our two inequality corrections.[21]

In the left panel of Figure 4, we plot the cumulative distribution of income for the year 2007, and for comparison we also plot the best lognormal fit of the distribution. As can be seen, the empirical distribution of income is pretty well approximated by a lognormal distribution. Nevertheless, the right panel of Figure 4 demonstrates that the lognormal fit is really poor for

[21]We could have in principle obtained disposable income by using the NBER TAXSIM program which calculates federal and state income tax liabilities from market income data. Nevertheless, this would have missed government transfers which are essential for understanding why disposable income is higher than market income for low-income individuals.

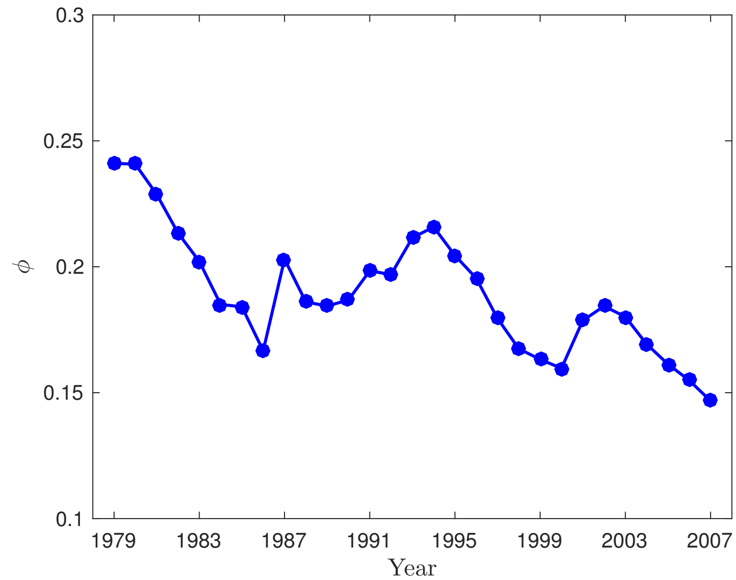

Figure 3: Progressivity Over Time

relatively high incomes, and in that range, a Pareto distribution appears to fit the data much better. More specifically, following Diamond and Saez (2011), this right panel plots the ratio $r_m/(r_m - r)$, with $r_m = \mathbb{E}(r_\varphi \mid r_\varphi > r)$, for different values of income r. Consistently with the properties of a Pareto distribution, for large enough income levels this ratio is relatively flat (at a value close to 1.5), whereas a lognormal distribution would predict this term to rise with income.

Having discussed the calibration of the progressivity parameter ϕ and the income distribution for each year in 1979-2007, we are left with the parameters ε, β and ρ. The size of the elasticity of taxable income ε has been the subject of heated debates in the academic literature. The influential work of Chetty (2012) has demonstrated, however, that when interpreting the wide range of estimated elasticities through the lens of a model in which agents face optimization frictions, an elasticity of taxable income to changes in marginal tax rates of around 0.5 can rationalize the conflicting findings of previous studies. With that in mind, we shall set $\varepsilon = 0.5$ in our benchmark calibration.[22]

Moving on to the substitutability parameter β, in our benchmark calibration we will set $\beta = 0.8$. The resulting elasticity of substitution $1/(1-\beta) = 5$ is slightly larger than that the one typically estimated with product-level trade (see Broda and Weinstein, 2006) or with firm-level mark-up data (see Bernard, Eaton, Jensen, and Kortum, 2003 or Antràs, Fort, and Tintelnot,

[22] It should be noted that five of the fifteen studies Chetty (2012) builds on to provide bounds on the intensive margin labor supply elasticity are based on the response of hours worked (rather than taxable income) to changes in marginal tax rates. In our model, these two elasticities are not identical due to the imperfect substitutability in tasks. We have however replicated Chetty's (2012) calculations restricting the analysis to the ten papers estimating taxable income elasticities. The resulting *intensive* margin elasticity is equal to 0.33, which is identical to the one obtained by Chetty (2012) when using all fifteen papers. Chetty (2012) also finds that the compensated and uncompensated elasticities taxable income are very similar, which helps to motivate our assumption of GHH preferences in (16). It is important to mention however that the evidence suggests that these elasticities appear to be higher for rich individuals than for poor ones, a feature that is absent in our model.

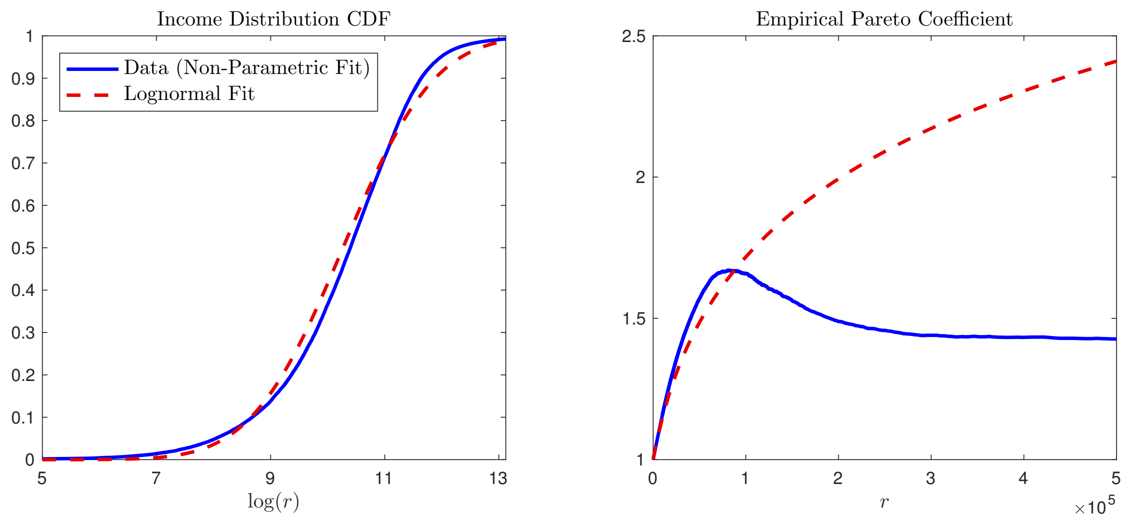

Figure 4: U.S. distribution of reported income (2007)

2014), but it seems reasonable to us to postulate that workers' tasks are more substitutable than the products that embody those tasks. When we quantitatively evaluate the effects of trade opening we will consider the sensitivity of our results to different values of β.

Finally, we discuss the calibration of the coefficient of inequality (or risk) aversion ρ. The often-used logarithmic utility case, which corresponds to $\rho = 1$, will provide a focal point for our quantitative analysis, but we readily admit that little is known about this parameter (especially when interpreted in terms of inequality aversion), and thus we will report results for various values of ρ ranging from $\rho = 0$ (no inequality aversion) all the way to $\rho = 2$.[23]

4.2 Evolution of the Inequality Correction Terms

Figure 5 depicts the evolution of the welfarist correction Δ and costly redistribution correction $(1 + \varepsilon\phi)\Theta^{\kappa}$ over the period 1979-2007 for the case $\rho = 1$. The smallest dot corresponds to the 1979 value of these terms, while the the largest dot corresponds to their 2007 value (the size of the dots grows over time). This graph embodies different pieces of information. Notice first that the welfarist discount factor Δ has been falling steadily over time, starting at a value of 0.757 in 1979 but ending at 0.587 in 2007. This decline necessarily reflects an increase in inequality in the distribution of disposable income. The graph however also shows that the causes of this increased dispersion in disposable income are twofold. On the one hand, the degree of tax progressivity has declined over time, something which was made clear in Figure 3, but which is also reflected by a noticeable upwards shift in the costly redistribution correction, which increased from 0.897

[23] Layard, Mayraz, and Nickell (2008) relate ρ to the degree to which marginal utility of income falls with income, and use survey data to argue that a value of $\rho = 1.26$ best explains the data. We have explored which value of ρ would rationalize a given year's observed degree of tax progressivity as being optimal in light of the social welfare function (22), and we have found the implied ρ to be much lower (between 0.35 and 0.5), which can be interpreted as reflecting a lower degree of inequality aversion or a higher influence of rich individuals in the setting of tax policies. This is quantitatively consistent with the findings of low inequality aversion for the United States, in contrast with Europe, in the macro-public-finance literature (see e.g. Chang, Chang, and Kim, 2016, and references therein).

in 1979 to 0.926 in 2007. If that was the only change in the environment, however, we would have expected the dots to line up along a negatively sloped locus. Instead, it is clear that the dots have also shifted *inwards* during this period, which necessarily implies an increase in the primitive determinants of inequality. In our closed-economy model, and holding the parameters β and ε constant, such an increase can only be generated by an increase in the dispersion of the distribution of ability. In section 5 we will show, however, that trade integration can generate an analogous inward shift even when the distribution of ability is held constant.

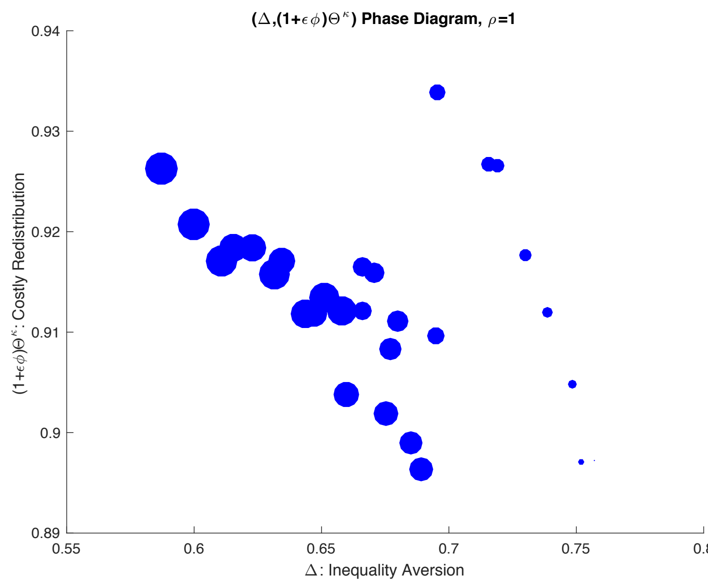

Figure 5: Evolution of the Welfarist and Costly-Redistribution Corrections

Table 1 further illustrates the consequences of these shifts for the evolution of U.S. social welfare over the period 1979-2007. The table uses equation (24) to decomposes changes in social welfare according to

$$\frac{W'}{W} = \frac{\Delta'}{\Delta} \times \frac{(1+\varepsilon\phi)'}{(1+\varepsilon\phi)} \times \underbrace{\frac{\Theta'^{\kappa}}{\Theta^{\kappa}} \times \frac{\tilde{W}'}{\tilde{W}}}_{Q'/Q},$$

where as indicated, changes in real income correspond to the changes in the product of Θ^{κ} and $\tilde{W}$ (for a constant ε and g over time).[24] The top panel of the Table performs the decomposition for various values of ρ. This serves to isolate the role of the welfarist correction in shaping the evolution of social welfare given the observed growth in real income. According to our data, mean real income grew 43.9% over 1979-2007, which translates into an annualized growth rate of 1.31% per year. In the absence of inequality aversion (or if the welfarist correction had not changed over time), the associated increase in U.S. social welfare over this period would have been slightly lower (37.9% or 1.15% per year) due to the increase in labor supply triggered by the decline in progressivity. Nevertheless, given that Δ fell considerably over time, the growth in social welfare was necessarily lower, and more so the higher is the degree of inequality aversion ρ.

[24]The share of government consumption in total GDP has indeed been relatively flat over the period 1979-2009, equalling 15.4% in 1979 and 15.3% in 2007.

The first column of Table 1 provides the discount factor by which the gross cumulated inequality-neutral growth rate 1.379 needs to be deflated in order to obtain the inequality-adjusted growth in social welfare for different values of ρ. The adjustment is potentially sizeable. For instance, for the logarithmic case ($\rho = 1$), the implied annual growth rate in social welfare is down to 0.24% (i.e., $(1.070)^{1/28} - 1$). Adopting a constant inequality aversion of 2, would actually result in a sizable *decline* of social welfare of 4.23% per year.

Table 1: Welfare, Inequality, and Costly Redistribution

	Change between 1979-2007				
$\rho =$	Δ	$1+\varepsilon\phi$	Θ^{κ}	$\tilde{W}$	W
0	1.000	0.958	1.079	1.334	1.379
0.25	0.923	0.958	1.079	1.334	1.273
0.5	0.868	0.958	1.079	1.334	1.196
1	0.776	0.958	1.079	1.334	1.070
1.5	0.637	0.958	1.079	1.334	0.879
2	0.216	0.958	1.079	1.334	0.298

Turning to the costly redistribution correction, Figure 5 indicates that $(1 + \varepsilon\phi)\Theta^{\kappa}$ has been rising over time – thus resulting in lower taxation inefficiencies – despite the observed increase in inequality. The reason for this is the marked decline in tax progressivity observed over these years. On account of this costly redistribution channel, social welfare has thus been growing by *more* than it would have in an economy without costly redistribution. More precisely, in a hypothetical *Kaldor-Hicks* economy with access to costless redistribution, average income would have grown by 1.03% (i.e., $(1.334)^{1/28} - 1$) per year on average, rather than the observed 1.31% annual growth.

4.3 Counterfactuals

Some readers might be struggling to wrap their heads around the interpretation of the costly redistribution correction since it involves a comparison of actual data with a hypothetical economy having access to costless redistribution. The usefulness of the adjustment will perhaps become more apparent when considering a couple of counterfactual exercises. As mentioned above, part of the reason why the welfarist term Δ decreased so markedly over time is the fact that U.S. redistribution became much less progressive over that period. One might then wonder: by how much would real disposable income and social welfare have increased if the degree of tax progressivity had been held constant at its 1979 level? And by how much would they have changed if tax progressivity had *increased* to ensure that the Atkinson measure of inequality (or our welfarist correction Δ) had not changed over 1979-2007?

Figure 6 provides answers to these questions for the benchmark case of $\rho = 1$ and also for a slightly lower value $\rho = 0.5$. The figure indicates that real disposable income would have grown at an average annual rate of 0.85% (instead of the observed 1.31%) if tax progressivity had been

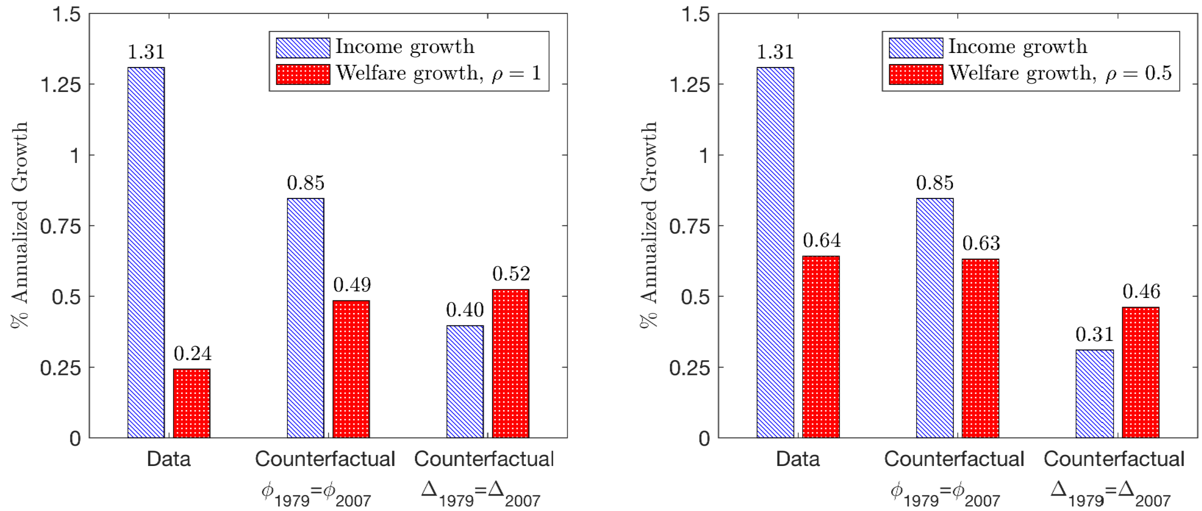

Figure 6: Counterfactuals: Holding ϕ constant and holding Δ constant

held constant at its 1979 level, while it would have grown at a significantly lower annual rate of 0.40% for $\rho = 1$ and 0.31% for $\rho = 0.5$ if tax progressivity had been raised so as to keep the Atkinson index constant at its 1979 level.[25] Despite the negative effect of these counterfactual policies on real income growth, the figure also shows that for $\rho = 1$, these policies would have increased social welfare by a nontrivial amount (from 0.24% to 0.49% and 0.63%, respectively). Conversely, for $\rho = 0.5$, social welfare would have instead declined had these policies been put in place.

5 Trade, Inequality and Costly Redistribution

For the remainder of this paper, we turn to the study of the interplay between inequality and costly redistribution in shaping the welfare consequences of trade liberalization episodes. With that goal in mind, in this section we consider a simple extension of our constant-elasticity model in section 3. Our model is highly stylized but generates trade-induced inequality via an intuitive mechanism that features prominently in the recent international trade literature, which builds on Melitz (2003). More specifically, our model captures the notion that agents can market their labor services in foreign markets only by incurring certain costs that are (at least in part) fixed in nature. Due to these costs, exporting is worthwhile *only* for the most productive agents in society. As a result, an even though all agents benefit as consumers from access to a larger measure of imperfectly substitutable tasks, trade integration raises real income disproportionately more for the highest-ability agents in society, thereby increasing income inequality.

These results are reminiscent of those delivered by models of trade-induced inequality featuring either labor market imperfections or matching and sorting of heterogenous agents. Relative

[25] In the second counterfactual, the effect is more pronounced for $\rho = 0.5$ than for $\rho = 1$ because the sensitivity of Δ to changes in ϕ is lower, the lower is ρ, and thus the necessary increase in ϕ is higher for $\rho = 0.5$ than for $\rho = 1$.

to these alternative models, the simplicity of our framework allows us to tractably incorporate behavioral responses to taxation into the study of the welfare effects of trade integration. Furthermore, as we shall see in section 6, our parsimonious model is particularly amenable to calibration and quantification.

5.1 A Simple Model of Trade-Induced Inequality

Consider a world economy consisting of $N+1$ symmetric regions analogous to the closed economy described in section 3. The fact that all regions are symmetric is essential for tractability because we lack the necessary cross-country data to discipline a calibration exercise in a world with asymmetric ability distributions, redistribution systems, and trade frictions. At the same time, it is not clear how the symmetric nature of our world economy will bias the quantitative results described below.[26]

As in our closed-economy model, individuals worldwide share the same preferences in (16) defined over the consumption of an aggregate final good and leisure, and they produce units of their differentiated task according to a linear technology in their labor effort. We assume that the aggregate final goods produced in different regions are perfect substitutes, and hence, given our symmetry assumption, they are not traded across regions. Conversely, all task (or intermediate inputs) produced worldwide are imperfectly substitutable and thus trade integration allows the final good to be produced more efficiently by combining a greater diversity of tasks provided by agents worldwide.

Agents can market their task in the local market at no cost, while in order to send the output of their task to other markets they need to incur trade costs which are both fixed and variable in nature. Specifically, in order to access $M \leq N$ foreign markets, any individual needs to pay M separate fixed costs $f(1)$, $f(2)$,..., $f(M)$. We interpret these fixed costs as being associated with the human capital investments necessary to make a worker's task marketable in a foreign country, and we characterize the fixed cost associated with the n-th market with the constant-elasticity function:

$$f(n) = f_x n^{\alpha}, \qquad \alpha \geq 0,\ n > 1. \tag{25}$$

Notice that f_x governs the average level of these fixed costs, while the parameter α shapes the curvature of the fixed cost function with respect to the number of foreign markets serviced. We introduce this parameter to allow us to more flexibly match a rich extensive margin of trade by which relatively more able individuals market their tasks in a larger number of markets (i.e., n_{φ} will be nondecreasing in φ). It should be stressed, however, that even when $\alpha = 0$, and

[26] Furthermore, the regions need not be interpreted as countries, but rather as trading blocks chosen to be symmetric with regards to the model's primitives. Our focus on N symmetric regions emphasizes the within-sector wage inequality mechanism, omitting the component of inequality that is likely to emerge from sectoral differences across countries (e.g., the Heckscher-Ohlin mechanism). Helpman, Itskhoki, Muendler, and Redding (2016) provide empirical evidence for the relevance of such focus, and Burstein and Vogel (2016) study quantitatively a related mechanism in a model which features both within- and between-sector forces, and find that the within-sector forces dominate and end up shaping the response of inequality to trade opening.

unlike in the Melitz (2003) model, more able individuals will still select into a (weakly) larger number of foreign markets. The reason for this is that, as we shall see, export entry decisions are interdependent in our framework.[27] On top of these fixed costs, when exporting to particular market, an agent needs to ship $\tau > 1$ units of task services for one unit to reach that foreign market. As a result, the export revenues obtained by an individual with ability φ from any foreign market j are given by $Q^{1-\beta}(q_\varphi/\tau)^\beta$, where q_φ is the number of units of task services shipped by that agent to that market.

An agent with ability φ thus invests in access to n_φ foreign markets, and optimally allocates the total output of its task y_φ across the markets, which yields a total revenue of:[28]

$$r_\varphi = \Upsilon_{n_\varphi}^{1-\beta} Q^{1-\beta} y_\varphi^\beta, \tag{26}$$

where

$$\Upsilon_{n_\varphi} = 1 + n_\varphi \tau^{-\frac{\beta}{1-\beta}}. \tag{27}$$

For a given output level y_φ, revenues are higher when sales are spread over a large number of markets n_φ because marginal revenue in each market falls by less relative to a situation in which all the output is sold in a single market, as in our closed-economy model. Given the symmetry in our model, we assume that individuals are indifferent with regards to which particular markets to serve, and choose to access a random subset of n_φ markets out of the total N foreign markets, thus maintaining symmetry across markets.

Market revenue is taxed according to a schedule $T(r)$ given by (11). Note that the tax is conditional only on market revenue, but not on the number of non-local markets served, n_φ. In other words, we assume that neither the ability of individuals nor their export investment decisions are observable. As in our constant-elasticity model in section 3, we continue to adopt the log-linear tax schedule introduced in section 2 and empirically motivated in section 4, so after-tax income for an agent with ability φ is given by

$$r_\varphi - T\big(r_\varphi\big) = k r_\varphi^{1-\phi},$$

where k is chosen to ensure balanced government budget, $\int_0^1 T\big(r_\varphi\big)\mathrm{d}H_\varphi = gQ$. Agents consume their after-tax income net of the fixed cost of entry,

$$c_\varphi = k r_\varphi^{1-\phi} - f_x \sum_{n=1}^{n_\varphi} n^\alpha, \tag{28}$$

and choose their labor supply ℓ_φ and export entry decisions n_φ to maximize utility (16) given

[27] We have also experimented with a variant of the model featuring heterogeneity of fixed costs of exporting across individuals, in a manner analogous to Eaton, Kortum, and Kramarz (2011) or Helpman, Itskhoki, Muendler, and Redding (2016).

[28] Formally, $r_n(y) = \max_{\{q_0, q_1, \ldots, q_n\}} \Big\{ Q^{1-\beta}\Big[q_0^\beta + \sum_{j=1}^n \big(q_j/\tau\big)^\beta\Big] \;\Big|\; \text{s.t. } \sum_{j=0}^n q_j = y \Big\}$, with the solution given by $q_0 = y/\Upsilon_n$ and $q_j = \tau^{-\beta/(1-\beta)} y/\Upsilon_n$ for $j = 1, \ldots, n$. Lastly, the agent with ability φ optimally choose $n = n_\varphi$ and $y = y_\varphi$, as we describe below, and we denote with $r_\varphi = r_{n_\varphi}\big(y_\varphi\big)$.

the production technology $y_\varphi = \varphi \ell_\varphi$, the revenue function (26) and the budget constraint (28). Given the quasi-linearity of preferences, we could alternatively treat the fixed cost of exporting as a utility cost (e.g., effort rather than spending on human capital investments), with no bearing for the welfare results discussed below.

The tasks sold in each market are combined by competitive firms into final output and sold on the local market. In equilibrium, trade across regions is balanced and the government spends all its net tax revenue on the final good, and thus the total expenditure on the local final good equals the total revenues of individuals in the region, $Q = \int r_\varphi \mathrm{d}H_\varphi$, which using (26) can be rewritten as:

$$Q = \left(\int_0^1 \Upsilon_{n_\varphi}^{1-\beta} y_\varphi^\beta \mathrm{d}H_\varphi \right)^{1/\beta}. \tag{29}$$

This completes the description of the open-economy environment of our framework. We will use the model to study the effect of a reduction in trade costs τ on social welfare, taking into account the effects of trade liberalization on aggregate income but also on inequality. As in our previous closed-economy model, we will measure social welfare according to the constant-inequality aversion function in equation (22).

5.2 Open-Economy Equilibrium

Solving the individual labor supply and market access problem results in the following before-tax revenue and utility schedules

$$r_\varphi = \left(\Upsilon_{n_\varphi}\right)^{\frac{(1+\varepsilon)(1-\beta)}{1+\varepsilon\phi}} r_{0\varphi}, \tag{30}$$

$$u_\varphi = \left(\Upsilon_{n_\varphi}\right)^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} u_{0\varphi} - f_x \sum_{n=1}^{n_\varphi} n^\alpha, \tag{31}$$

where

$$r_{0\varphi} = \left(\beta(1-\phi)k\right)^{\frac{\varepsilon}{1+\varepsilon\phi}} \left[Q^{1-\beta}\varphi^\beta\right]^{\frac{1+\varepsilon}{1+\varepsilon\phi}} \quad \text{and} \tag{32}$$

$$u_{0\varphi} = \frac{1+\varepsilon\phi}{1+\varepsilon} k \left(r_{0\varphi}\right)^{1-\phi} \tag{33}$$

correspond to before-tax revenue and utility of an agent with ability φ that only sells locally, and where $\Upsilon_{n_\varphi} = 1 + n_\varphi \tau^{-\frac{\beta}{1-\beta}}$ with $n_\varphi \in \{0, 1, ..., N\}$ being determined by

$$\left[\Upsilon_{n_\varphi}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \Upsilon_{n_\varphi-1}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}}\right] u_{0\varphi} \geq f_x n_\varphi^\alpha, \tag{34}$$

$$\left[\Upsilon_{n_\varphi+1}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \Upsilon_{n_\varphi}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}}\right] u_{0\varphi} < f_x (n_\varphi+1)^\alpha. \tag{35}$$

Because $(1+\varepsilon)(1-\beta) < 1$ and $\alpha \geq 0$, the existence of a unique $n_\varphi \in \{0, 1, ..., N\}$ is guaranteed.

These equilibrium expressions are cumbersome, but they can be used to establish that, starting from autarky, trade integration induces an increase inequality in our framework. To see this, define an *increase in inequality* in a given variable x as a situation in which the ratio $x_{\varphi^H}/x_{\varphi^L}$ for two individuals with abilities $\varphi^H > \varphi^L$ is either left unchanged or increased, with this ratio being increased for at least a pair of individuals. With this definition in hand, we can then show (see Appendix A.4) the following result, which is reminiscent of the main result in Helpman, Itskhoki, and Redding (2010):

Proposition 3. *A move from autarky to a trade equilibrium in which some (but not all) individuals export to some markets necessarily increases inequality in pre-tax and after-tax real income and in utility levels.*

The intuition for the result is simple. In the presence of fixed costs of exporting, some relatively low-ability individuals will not be able to profitably market their task in foreign markets, while these same agents will now face increased competition from foreign high-ability individuals selling their task in their local market.

This result is illustrated in Figure 7 which plots the Gini coefficient and coefficient of variation of real market and disposable income for different levels of trade costs. As is clear, these measures of inequality are minimized for the largest values of variable trade costs τ, which indeed place the economy close to autarky. The figure also shows that the effect of trade integration on inequality need not be monotonic. In fact, it is straightforward to show that if fixed costs of exporting are sufficiently low, a reduction in iceberg trade costs that leads all individuals to market their tasks in all regions will necessarily reduce inequality. This is because the level of inequality associated with an economy in which all individuals sell in all markets is identical to the level of inequality under autarky. Despite the fact that trade cost reductions could *theoretically* reduce income inequality, our calibration exercise (to be discussed in detail in the next section) indicates that it would take a significant decline in trade costs relative to those estimated for the U.S. in both 1979 and 2007 to enter the region in which trade is associated with reduced rather than increased income inequality (see the dashed vertical lines in Figure 7).

So far we have focused on the implications of the model for the effect of trade on inequality and we have been able to state Proposition 3 without solving for the endogenous aggregate variables k and Q. In order to study the welfare gains from trade integration it is, however necessary to solve for these objects. Aggregate income Q can be solved as a function of k as the fixed point of

$$Q = \int r_{\varphi}(Q,k)\,\mathrm{d}H_{\varphi},$$

where $r_{\varphi}(Q,k)$ is obtained by combining equations (30)-(35). The value of Q and k can then be obtained by noting that $k = (1-g)\int r_{\varphi}(Q,k)\,\mathrm{d}H_{\varphi}/\int (r_{\varphi}(Q,k))^{1-\phi}\,\mathrm{d}H_{\varphi}$ as in equation (18).

Manipulating these equations, we show in Appendix A.4 that:

Proposition 4. *A move from autarky to a trade equilibrium in which some individuals export to some markets necessarily increases aggregate real income Q. Furthermore, the utility of all agents is higher than under autarky.*

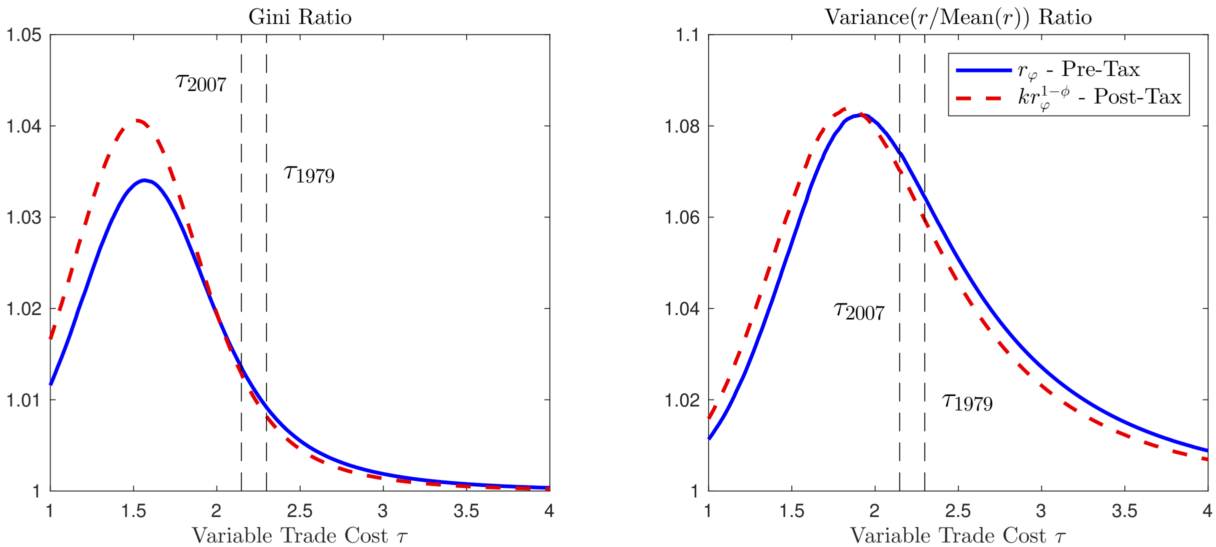

Figure 7: Trade Integration and Income Inequality

In sum, despite the fact that trade typically decreases the relative revenues obtained by low-ability individuals, the reduction in the price index faced by these individuals when acting as consumers is always large enough to leave them at least as well off as before the reduction in trade costs. Despite the existence of Pareto gains from trade, the fact that some agents benefit more than others will have significant implications for the quantifying of the overall social welfare gains from trade, as we shall see in section 6.

5.3 Social Welfare and the Gains from Trade

Once we have solved for the aggregates of the model, we can plug them back into (31) and (33) and invoke (22) to compute social welfare in the open-economy equilibrium. As in section 3, we can denote by $\tilde{W}_T$ the social welfare in the economy with zero tax progressivity ($\phi = 0$) and no inequality aversion ($\rho = 0$), and use this definition to decompose social welfare as

$$W = \frac{\left[\mathbb{E}\left(u_\varphi\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}}{\mathbb{E}u_\varphi} \times \frac{\mathbb{E}u_\varphi}{\tilde{W}} \times \tilde{W} = \Delta_T \times \Theta_T \times \tilde{W}_T. \tag{36}$$

Welfare is thus the product of the potential welfare level $\tilde{W}_T$ attainable in the absence of tax progressivity or inequality aversion, and two terms, Δ_T and Θ_T, that are analogous to the welfarist and costly-redistribution corrections developed in sections 2 and 3.

Given the equilibrium values of u_φ in (31), the welfarist correction term Δ_T can easily be computed for a particular value of ρ. Furthermore, the fact that, by Proposition 3, trade integration increases inequality in utility levels implies that (see Appendix A.4):

Proposition 5. *Relative to its value under autarky, Δ_T is strictly lower in a trade equilibrium in which some (but not all) individuals export to some markets. Furthermore, the welfare gains from trade are strictly decreasing in the degree of inequality aversion ρ.*

This result formalizes the fact that in the presence of trade-induced inequality, the 'welfarist' gains from trade will necessarily be lower than those implied by the Kaldor-Hicks criterion, and more so, the more society is averse to inequality.

How is the costly-redistribution correction Θ affected by trade? The increase in inequality demonstrated in Proposition 3 would appear to hint at a reduction in Θ following a move from autarky to some form of trade integration. Nevertheless, the result does not hold generally because, in a trade equilibrium, redistribution policy not only reduces the incentives to supply labor given a trade status, but also shapes the extensive margin decisions of agents as to whether service particular foreign markets.[29] For certain parameter values, an open economy can be closer to its costless redistribution counterfactual than its autarky counterpart. Despite this theoretical ambiguity, in our quantitative analysis with realistic parameter values (see section 6), we have found that Θ is always reduced when moving from autarky to an equilibrium with positive trade flows. This of course implies that our costly redistribution correction will also modify the magnitude of the gains from trade *downwards*.

6 Calibration and Trade Counterfactuals

We are now ready to turn to a quantitative exploration of our model centered on the U.S. experience over the period 1979-2007. Our ultimate goal is to quantify the role of trade-induced inequality in shaping the welfare consequences of the observed rise in trade integration over the period 1979-2007. We proceed in two steps. On a first pass, we calibrate our model to match certain key moments of the 2007 United States economy. We then increase trade frictions to bring the openness of the U.S. economy back to its 1979 level (and also back to autarky, in an auxiliary exercise). This allows us to compute the effect of changes in trade openness on aggregate income and on inequality, thereby allowing us to gauge the quantitative importance of the two corrections developed in this paper. More specifically, we seek to answer the following questions: how large are the gains from trade for different degrees of inequality aversion? How large would the gains from trade have been if costless redistribution had been available?

6.1 Calibration

The calibration of our model to 2007 U.S. data is analogous in many ways to the one we performed in section 4 for the closed-economy version of the model during the period 1979-2007. In particular, we continue to set $\beta = 0.8$ and $\varepsilon = 0.5$, while we again back out the tax progressivity parameter ϕ by regressing the logarithm of CBO post-tax and transfer income on the logarithm of market income, though in this case we focus on the year 2007. As in the different panels of Figure 2, the fit of this simple log-linear regression is equally remarkable in that year and delivers an estimate of $\phi = 0.147$ with an R-squared of 0.995.

The only new sets of parameters to calibrate in the open economy are (i) the number of

[29]This feature of the model bears some resemblance to analysis of optimal income taxation with both intensive and extensive margins of labor supply responses, as in Saez (2002).

symmetric foreign regions N, (ii) the iceberg cost parameter τ, and (iii) the parameters f_x and α determining the structure of fixed costs of exporting. According the World Bank's world development indicators, the U.S. accounted for 18.3% of world (PPP-adjusted) GDP in 2007, so we set $N = 5$ in our benchmark calibration, though we will also present some sensitivity results for different values of N in section 6.3.

The calibration of the other trade parameters is more involved. Realizing that the income distribution produced by the model is crucially affected by the exporting decisions of agents, we jointly calibrate the ability distribution H_φ and the trade parameters (τ, f_x, α) to exactly match the 2007 distribution of market income (from the public use samples of U.S. Federal Individual Income Tax returns) as well as three moments of the U.S. trade sector. These three "trade moments" are: (M1) the U.S. trade share in 2007, defined as the ratio of the average of U.S. exports and imports to gross output (7.74%); (M2) the share of exporters' sales in the total sales of U.S. firms (61.8%); and (M3) the share of U.S. exports accounted by for exporters that sell to more than five foreign markets (88.9%).

The choice of these three moments is motivated by the following two considerations. First, it is not clear what a firm is in our model, and thus we unfortunately cannot rely on the large number of firm-level moments developed in the literature (e.g., Bernard, Jensen, Redding, and Schott, 2007) to discipline our model. Instead, we rely on moments aggregated across firms (and workers). Second, we target moments that we think are particularly useful in jointly identifying the parameters of interest. In that respect, (M2) and (M3) both relate to the extensive margin of trade, which is crucially affected by the shape of the fixed cost function in (25). More specifically, (M3) is a measure of concentration that should be largely shaped by the curvature α of the fixed cost function. The overall trade share (M1) is determined by a combination of the extensive and intensive margins of trade, and thus it seems reasonable that, conditional on the other two moments, this moment will help pin down the empirically relevant value for iceberg trade costs.

We provide more details on the sources of data used to compute these moments in Appendix A.5, where we also include a discussion of the technical aspects of the calibration. The resulting parameter estimates are $\tau = 2.147$, $f_x = 675$, and $\alpha = 0.554$. Our estimated iceberg trade costs may appear to be rather high, but it is important to emphasize that we are calibrating the model to the entire U.S. economy, rather than to its manufacturing sector, as is standard in quantitative models of trade.[30] Conversely, our estimated fixed costs of exporting might appear low relative to those estimated in the literature (e.g., Das, Roberts, and Tybout 2007), but note that ours apply at the individual level, while the literature has estimated them at the firm level. Furthermore, in our framework, these fixed costs rise with the number of markets serviced.

When performing our counterfactuals, we hold all parameters fixed, including the distribution of ability, and we first set $\tau_{1979} = 2.298$ to match the 1979 trade share of 4.90% in the data. This amounts to rolling back the U.S. economy to its openness level in 1979, so we can isolate the effects of trade integration on aggregate income and inequality. In a second more extreme

[30] When calibrating our model to the U.S. manufacturing sector we indeed back out a smaller value of τ ($\tau = 1.79$), which is very much in line with those in Anderson and van Wincoop (2004) and Melitz and Redding (2015).

counterfactual, we study a shift to autarky by setting $\tau_{autarky} = +\infty$, while again holding all other parameters fixed at their 2007 level.

6.2 Counterfactuals

Table 2 reports the implications of a move to the 1979 openness level and to autarky for aggregate consumption, for aggregate welfare in the absence of inequality aversion (i.e., 'Kaldor-Hicks welfare'), and for the Gini coefficient of disposable income. For our benchmark taxable income elasticity of 0.5, a move to 1979 openness levels would reduce consumption by 1.2% percent and Kaldor-Hicks welfare by 1.1%, the difference reflecting a lower labor supply (and thus higher leisure) in a world with depressed export sales. A move to autarky naturally magnifies these numbers, which are (respectively) 3.4% and 3.2% percent in that case.[31] Importantly for our purposes, the last two columns of Table 2 show that these income and Kaldor-Hicks welfare losses are accompanied by nonnegligible declines in inequality, with the Gini coefficient falling by 0.5% in the 1979 counterfactual and by 1.3% in the autarky counterfactual. These numbers are somewhat dwarfed by the actual increases in the Gini coefficient observed during the period 1979-2007 (see Figure 1), but of course many other forces were at play in the last few decades contributing to the drastic increase in income inequality. As we will next see, even when trade might have been a small contributor to the observed increase in inequality, such a contribution still has nontrivial consequences for the measurement of the welfare gains from trade.

Table 2 also shows that the real consumption gains from trade are higher, the higher is the taxable income elasticity ε. This is consistent with the findings of Arkolakis and Esposito (2014). Notice, however, that the amount of inequality induced by trade opening also increases with ε.

Table 2: Welfare Gains from Trade and Induced Inequality

	% Consumption Gains		% K-H Welfare Gains ($\rho = 0$)		% Increase in Gini	
	τ_{1979}	$\tau = \infty$	τ_{1979}	$\tau = \infty$	τ_{1979}	$\tau = \infty$
$\varepsilon = 0.25$	0.8	2.4	0.8	2.3	0.4	1.1
$\varepsilon = 0.5$	1.2	3.4	1.1	3.2	0.5	1.3
$\varepsilon = 1$	2.0	6.0	1.9	5.6	0.6	1.6

We are now ready to put the tools developed in this paper to use, and invoke equation (36) to compute modified social welfare gains from trade that take into account the inequality-enhancing consequences of trade opening. In order to better compare the impact of inequality on the gains from trade under different counterfactual exercises, parameter values, and calibration approaches, we will focus on reporting what we refer to as the *welfarist modified statistic* and the *costly redistribution modified statistic*. The former, which we denote by Δ^{Stat}, corresponds to the factor by which the social welfare gains from trade in the absence of inequality aversion need to be multiplied to obtain the gains from trade whenever $\rho > 0$. The latter, denoted by Θ^{Stat}, is instead the factor by which the gains from trade in a hypothetical Kaldor-Hicks economy need

[31] These numbers are broadly in line with those obtained in the literature calibrating models of trade featuring no income dispersion within countries (Costinot and Rodríguez-Clare, 2013).

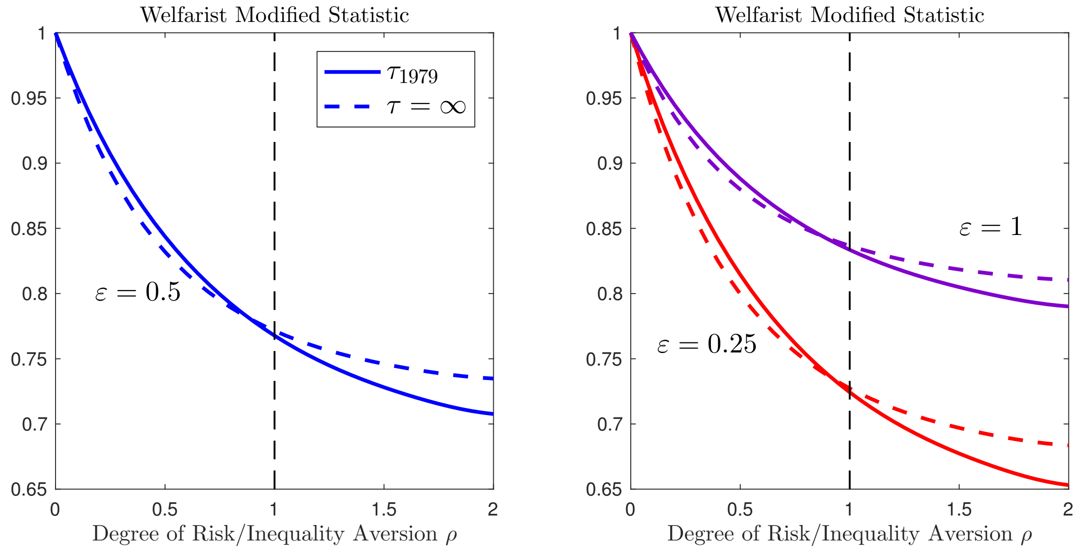

Figure 8: Welfarist Adjustment to the Welfare Gains from Trade

to be multiplied in order to obtain the gains from trade in an inequality-neutral economy in which redistribution is costly. In more formal terms, and analogously to the decomposition in equation (36), we are effectively decomposing the (net) gains from trade as follows:[32]

$$\frac{W'}{W} - 1 = \underbrace{\left(\frac{\frac{W'}{W} - 1}{\frac{W'/\Delta'_T}{W/\Delta_T} - 1} \right)}_{\Delta^{Stat}} \times \underbrace{\left(\frac{\frac{\tilde{W}'_T \cdot \Theta'_T}{\tilde{W}_T \cdot \Theta_T} - 1}{\frac{\tilde{W}'_T}{\tilde{W}_T} - 1} \right)}_{\Theta^{Stat}} \times \left(\frac{\tilde{W}'_T}{\tilde{W}_T} - 1 \right)$$

Given our theoretical results, we expect both Δ^{Stat} and Θ^{Stat} to be less than one, thus leading to a downwards adjustment to the gains from trade.

Figure 8 plots the welfarist modified statistic in each of the two counterfactuals for our benchmark parameter values, and for different values of ρ and ε. The left panel focuses on our benchmark with $\varepsilon = 0.5$, while the right panel replicates the figure for $\varepsilon = 0.25$ and $\varepsilon = 1$. Interestingly, even though the effects of trade on aggregate income and income inequality are quite different in the two counterfactuals under study (see Table 2), the figure indicates that the welfarist modified statistic turns out to be almost identical in the two counterfactuals (corresponding to solid and dashed lines respectively), regardless of the value of ε. For the case of logarithmic utility ($\rho = 1$) and a benchmark taxable income elasticity of $\varepsilon = 0.5$, we have that $\Delta^{Stat} = 0.77$ for both the 1979 and the autarky counterfactuals, which implies that inequality aversion is associated with welfare gains that are 23% lower than in a world with an inequality-neutral ($\rho = 0$) social planner. For $\varepsilon = 0.25$ and $\varepsilon = 1$, the analogous factors are 0.73 and 0.84, respectively, which are associated with 27% and 16% lower gains from trade.

Naturally, this welfarist modified statistic approaches 1 as ρ goes to 0, but notice that even for relatively low values of ρ, such as $\rho = 0.5$, the adjustment still erodes between 10% and 20%

[32]This decomposition applies since, according to (36), $W/\Delta_T = \tilde{W}_T \cdot \Theta_T$.

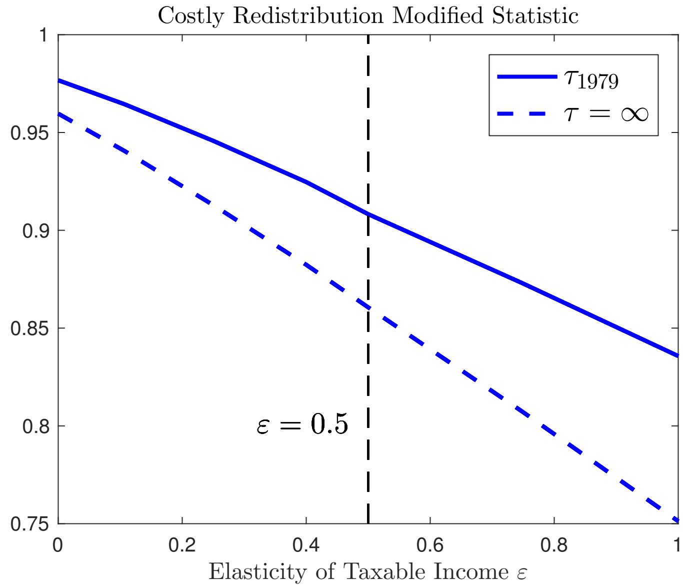

Figure 9: Costly Redistribution Adjustment to the Welfare Gains from Trade

of the gains from trade. On the other hand, and although the gains from trade decline in ρ (see Proposition 5), we find that the relationship between Δ^{Stat} and ρ quickly flattens. As a result, even for very large value of ρ, such as $\rho = 2$, our welfarist adjustment is unlikely to eliminate more than one-third of the gains from trade.

Turning to our costly redistribution modified statistic Θ^{Stat}, Figure 9 plots this statistic for the 1979 and autarky counterfactuals, and for various values of ε. Because the inefficiencies associated with costly redistribution are increasing in the elasticity of taxable income ε, it is not surprising that Θ^{Stat} is declining in ε. Unlike in the case of Δ^{Stat}, the specific counterfactual under consideration does appear to matter for the size of Θ^{Stat}, with the autarky counterfactual being associated with a larger downwards adjustment. For our benchmark value of $\varepsilon = 0.5$, we find that $\Theta^{Stat} = 0.86$ in our autarky counterfactual and $\Theta^{Stat} = 0.91$ in the 1979 counterfactual. This implies that in the presence of costless redistribution, the gains from trade would have been about 16% (i.e., $1/0.86 - 1$) higher in the autarky counterfactual and about 10% higher in the 1979 counterfactual.

6.3 Robustness

Although our benchmark calculations rely on parameter values and an estimation approach that we find trustworthy, we are well aware that a few of our choices are not uncontroversial, so it is important to explore the sensitivity of our results to alternative approaches and calibrations.

Above, we have already illustrated how our quantitative results vary with the degree of inequality aversion ρ and with the value of the taxable income elasticity ε (see Figures 8 and 9). The welfarist modified statistic Δ^{Stat} is naturally lower the higher is ρ, but the downward revision to the gains from trade remains moderate even for very large values of ρ (such as $\rho = 2$). Similarly, we have found that Δ^{Stat} tends to be higher and Θ^{Stat} tends to be lower,

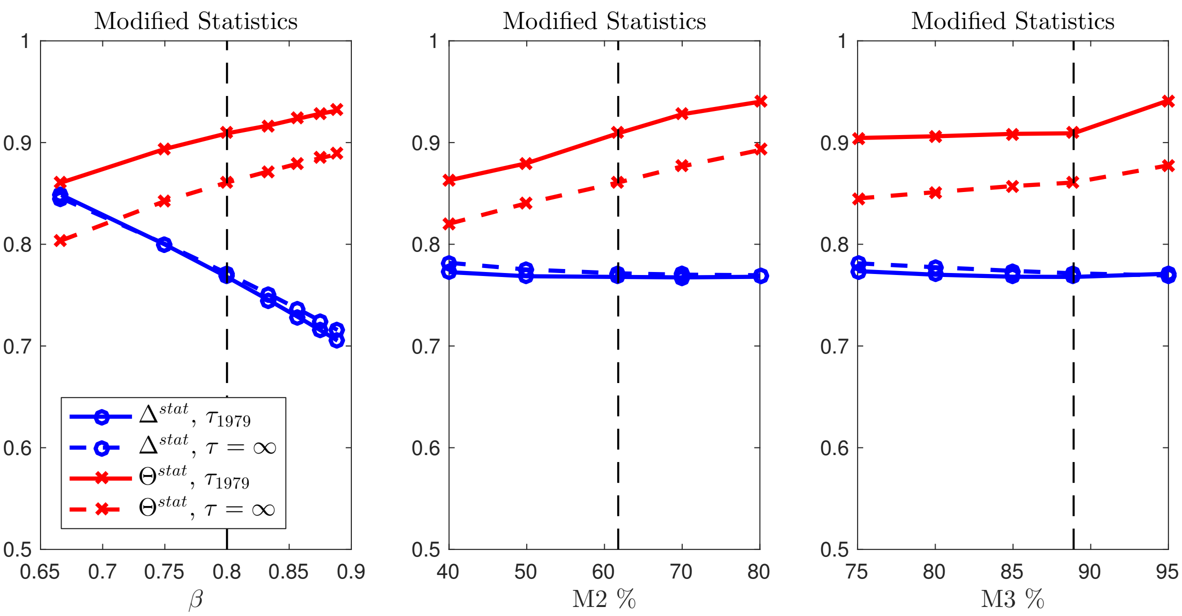

Figure 10: Robustness to β and to $M2$

the higher is ε, but these modified statistics remain in a relatively tight range even for very low ($\varepsilon = 0.25$) and very high ($\varepsilon = 1$) values of the taxable income elasticity.

We next turn to the substitutability parameter β, which we set to $\beta = 4/5 = 0.8$ in our benchmark calibration. We chose this value because it seems reasonable to assume a slightly higher elasticity of substitution across workers' tasks than across the products that embody those tasks, with the latter substitutability typically associated with a value of around $\beta = 3/4$. One might wonder, however, how our results would change if we instead chose a higher (or a lower) value of β. Intuitively, the level of β is important in determining the size of the gains from trade, but it is also crucial in shaping the extent of income inequality arising from our model. A key question is then: how does the trade-off between aggregate income growth and income inequality following trade liberalization vary with the value of β? In the left panel of Figure 10, we plot our two modified statistics Δ^{Stat} and Θ^{Stat} for various values of β and for both the 1979 and the autarky counterfactuals. In order to isolate the role of variation in β from that of variation of ε, in our calculations we adjust the value of γ so that ε remains at 0.5 regardless of the value of β. We also set $\rho = 1$ to keep the description of the results manageable. As Figure 10 indicates, changing β has a moderate effect on our estimates. The welfarist modified statistic, which was 0.77 in our benchmark calculations in both counterfactuals, is only reduced to 0.71–0.72 when β is increased to 8/9, while reducing β to 2/3 only raises Δ^{Stat} to 0.85. The effect of β on the costly redistribution modified statistic is similarly limited, with Θ^{Stat} remaining in the range 0.8–0.93. The reason for these modest changes is that although the gains from trade tend to fall monotonically with β, it is also the case that the effect of trade on inequality is lower, the higher is β.[33]

[33] It is worth noting that although the *change* in inequality following trade opening is lower, the higher is β, the *level* of inequality given a level of trade frictions is higher, the higher is β.

Consider next the degree of tax progressivity ϕ. The remarkably tight log-linear relationship between market income and after-tax income in the CBO data makes us confident about $\phi = 0.147$ capturing the degree of tax progressivity in 2007. It is less clear, however, than in performing our counterfactuals it is reasonable to focus solely on this parameter value. For instance, one might conjecture that the fact that the welfarist correction generally leads that a larger adjustment to the gains from trade than the costly redistribution correction is driven by the fact that, in 2007, tax progressivity was at its lowest level in the period 1979-2007 (see Figure 3). This conjecture can be evaluated by repeating our calculations in our benchmark case but instead setting ϕ to its average value in 1979-2007, namely $\phi = 0.189$. We report the results in column (b) of Table 3. In such a case, in the autarky counterfactual, the welfarist and costly redistribution modified statistics turn out to be almost identical, and both suggest an adjustment of around 20% of the gains from trade.[34]

Beyond these considerations regarding the level of ϕ to use in the calculations, it might also seem questionable that we hold this parameter constant in our counterfactuals. An intriguing possibility is whether part of the decline in tax progressivity observed in the period 1979-2007 may be ascribed to the trade shock our counterfactual is trying to isolate. Indeed, and consistently with the results of Itskhoki (2008), in numerical simulations of our model, we find that the social welfare maximizing value of tax progressivity ϕ^* is typically lower in a trade equilibrium than under autarky.[35] In our benchmark calibration of the model, using the degree of inequality aversion ρ that makes $\phi = 0.147$ optimal in 2007, the counterfactual optimal degree of tax progressivity in 1979 and under autarky is indeed higher and around $\phi = 0.16$. Although the difference may seem small, it has a pretty dramatic effect on our modified statistics, as shown in column (c) of Table 3. The endogenous decline in tax progressivity magnifies trade-induced inequality in the distribution of disposable income, and the welfarist modified statistic Δ^{Stat} drops to 0.44 in the 1979 counterfactual and to 0.71 in the autarky counterfactual. This implies that the gains from trade may be half of those under an inequality-neutral social planner. Conversely, trade now endogenously reduces the distortions associated with costly redistribution, and as a result the costly-redistribution modified statistic becomes larger than one (and significantly so in the 1979 counterfactual), reflecting that the gains from trade may well be *higher* than in a hypothetical Kaldor-Hicks economy in which the (full) efficiency of redistribution is held constant before and after the trade shock.

In our benchmark results, we have assumed that the world is composed of the U.S. and $N = 5$ additional blocks that are symmetric to the U.S. in all respects. As explained above, it would be interesting to relax the symmetry assumption, but unfortunately this is not feasible given data constraints. On the other hand, our quantitative analysis can easily be performed

[34] We have also redone our calculations fixing the value of ϕ to its 1979 level, 0.241, which is the highest in the period 1979-2007. In that case, Δ^{Stat} is around 0.86 in both counterfactuals, while Θ^{Stat} equals 0.84 and 0.74 in the 1979 counterfactual and autarky counterfactuals, respectively. These results again illustrate the role of tax progressivity in shaping the relative size of our two adjustments to the gains from trade.

[35] Formally, ϕ^* is the level of tax progressivity that maximizes social welfare (36) for a given inequality aversion ρ and trade openness τ. We find that ϕ^* decreases with trade openness (i.e., as τ decreases) in our calibrated model. Intuitively, this is because the cost of taxation increases faster with openness than trade-induced inequality, given our welfare criterion.

Table 3: Other Sensitivity Tests

		Benchmark (a)	Avg. ϕ (b)	Endog. ϕ (c)	$N=3$ (d)	$N=7$ (e)	Manuf. (f)	$\varphi \sim \log \mathcal{N}$ (g)
Δ^{Stat}	1979	0.77	0.81	0.44	0.77	0.77	0.77	0.76
	Autarky	0.77	0.82	0.71	0.77	0.78	0.77	0.78
Θ^{Stat}	1979	0.91	0.88	1.81	0.93	0.90	0.99	0.93
	Autarky	0.86	0.81	1.04	0.88	0.85	0.94	0.83

for various values of N. As shown in columns (d) and (e) of Table 3, the welfarist and costly redistribution modified statistics are only marginally different when setting $N = 3$ or $N = 7$ (and recalibrating the model parameters to still fit the empirical moments M1-M3).

Turning to the trade cost parameters τ, f_x, and α, we have calibrated these — together with the ability distribution H_φ — to match the 2007 U.S. trade share, the share of exporters' sales in the total sales of U.S. firms, and the share of U.S. exports accounted by for exporters that sell to more than five foreign markets. Although, we borrow the second moment (which we refer to as M2) from Antràs, Fort, and Tintelnot (2014), one might be concerned about the assumptions underlying the construction of such a moment based on the aggregation of firm-level data, as well as about the mapping of this moment to our theoretical model in which workers (rather than firms) export. With that in mind, in the middle panel of Figure 10 we report the values of Δ^{Stat} and Θ^{Stat} for calibrations involving various alternative values of M2, ranging from 40% all the way to 80%. As the figure indicates, the welfarist modified statistic Δ^{Stat} is remarkably insensitive to M2 and remains in the range 0.76-0.77. The costly-redistribution statistic Θ^{Stat} is a bit more sensitive to this moment, but the percentage adjustment to the gains from trade remains in a fairly narrow range. Similarly, one might have concerns about our choice of the third moment (M3), which captures the concentration of export volumes among producers that sell in many markets. As in the case of M2, our quantitative results are however fairly insensitive to this moment. More precisely, the right panel of Figure 10 shows that, even when reducing M3 from our benchmark value of 88.9% down to 75%, or when increasing it up to 95%, Δ^{Stat} remains in the neighborhood of 0.77–0.78. The costly-redistribution statistic Θ^{Stat} is equally remarkably insensitive to reductions in M3, but it does slightly increase relative to our benchmark case when M3 is brought up to 95%.

Another distinctive aspect of our calibration, relative to the bulk of quantitative work in international trade, is that we have attempted to calibrate the *entire* U.S. economy, rather than just its manufacturing sector. This choice was driven by the fact that our key NBER-IRS pre-tax income distribution data applies to workers in all sectors of the U.S. economy, rather than just in manufacturing. Nevertheless, assuming that the income distribution of workers engaged in manufacturing is similar to that of workers in other sectors, we can easily repeat our calibration but target trade moments related to the manufacturing sector. In particular, we set the trade share (M1) in manufacturing to 23.1% (as reported by the U.S. Census Bureau), the second moment (M2) for manufacturing firms to 90.1% (as in Antràs, Fort, and Tintelnot, 2014), and the third moment (M3) also for manufacturing firms to 95.5% (as reported by the U.S. Census Bureau). The results are presented in column (f) of Table 3. As might have been guessed from our previous results, the welfarist modified statistic again turns out to be unaffected by this changes (and remains at 0.77 in both counterfactuals), while the costly-redistribution statistic is now larger and even approaches 1 in our 1979 counterfactual.[36]

Our final sensitivity tests relate to our calibration of the ability distribution. First, we have adopted a non-parametric approach in which the ability distribution is chosen to exactly match the resulting pre-tax income distribution from Federal IRS returns, but one may wonder whether our results would be significantly different if we followed a more parametric approach. For instance, the Public Finance literature has recently been concerned with the implications of the shape of the right tail of the income distribution for optimal income taxation (see, for instance, Diamond and Saez, 2011). With that in mind, and given that a lognormal distribution appears to match the U.S. income distribution rather well except for its far right tail, we next assess the robustness of our results to an alternative approach in which the distribution of ability is assumed to be lognormal. Quite intuitively, we find that such a parametric approach tends to underpredict both the welfarist and costly redistribution corrections Δ and Θ, since it tends to underpredict the mass of individuals at the far right of the income distribution. Despite these biases, as column (g) indicates, imposing a lognormal distribution only has a very modest effect on the modified statistics. The key for understanding this insensitivity is that these modified statistics are shaped by how the welfarist and costly-redistribution corrections change when trade frictions are changed, rather than by the level of these corrections.

Beyond the particular shape of the ability distribution, a valid concern is whether it is reasonable to hold the ability distribution constant in our counterfactual exercises. Indeed, if part of the productivity shifter φ is determined by human capital investments, given the long horizons that we consider, it seems plausible that the changes in marginal revenue induced by trade opening would affect these human capital investments. A full exploration of this possibility is beyond the scope of this paper. Nonetheless, we have developed a simple extension of our model with human capital investment that responds to a trade shock.[37] We show that this generalization is essentially isomorphic to the baseline model, but features a (long-run) elasticity of taxable income that is larger than the short-run one (ε). The resulting effects on the modified statistics Δ^{Stat} and Θ^{Stat} can then be read off our sensitivity analysis with respect to ε in Figures 8 and 9. If, for instance, the long-run elasticity is as high as 1, we have $\Delta^{Stat} = 0.84$ and $\Theta^{Stat} = 0.75$. For even higher values of ε, the welfarist modified statistic becomes even larger (0.88 for $\varepsilon = 1.5$ and 0.92 for $\varepsilon = 2$), while the costly-redistribution modified statistic

[36]The value of Θ^{Stat} close to one is tied to the fact than in our calibrated 2007 manufacturing sector, around 55% of agents export their tasks, while only 19% do so in the economy-wide calibration. As a result, the marginal exporters tend to face much lower marginal taxes in the manufacturing calibration than in the economy-wide one, and the export participation margin response to an increase in τ is thus much more pronounced in the latter case than in the former.

[37]In particular, we consider an individual worker production function $y = \varphi \ell^{1-a} h^{a}$, where ℓ is labor hours or labor effort, h is human capital and only adjusts in the long run, and a is the output elasticity with respect to human capital.

becomes significantly lower (0.64 for $\varepsilon = 1.5$ and 0.53 for $\varepsilon = 2$). Intuitively, as the elasticity of taxable income becomes larger and larger, costly redistribution becomes a larger burden to the realization of the real income gains from trade, but such a scenario also tends to keep in check the extent to which trade integration increases the dispersion in the distribution of disposable income, thus moderating the welfarist adjustment to the gains from trade.

7 Conclusions

In this paper, we have explored the welfare consequences of trade integration in an environment in which trade-induced inequality is partly mitigated by a progressive income tax-transfer system. Despite the progressive nature of taxation, trade integration leads to an increase in inequality in the distribution of disposable income. We have argued that, under these circumstances, the application of the Kaldor-Hicks criterion to quantitatively evaluate the welfare gains from trade is not devoid of value judgments. More specifically, unless one is willing to assume that a dollar in the hands of a poor individual has the same social value as a dollar in the hands of a rich individual, trade liberalization episodes that increase the real disposable income of some individuals but reduce that of others cannot be evaluated by simply adding those real incomes. Furthermore, in situations in which trade integration benefits some agents in society disproportionately, the progressivity of the tax system implies that these fortunate individuals will necessarily transition into higher marginal tax brackets, so they will naturally adjust their labor supply (or effort in production) in a way that diminishes the realized gains from trade relative to a situation in which redistribution was performed in a nondistortionary manner. In this paper, we have formalized these insights and we have developed welfarist and costly redistribution corrections to standard measures of the gains from trade integration. Under plausible parameter values, these corrections are nonnegligible: trade-induced increases in the dispersion of disposable income reduce the gains from trade by about 20%, and these welfare gains would be about 15% larger if redistribution was carried via non-distortionary means.

References

Aguiar, M., and M. Bils (2015): "Has consumption inequality mirrored income inequality?," *The American Economic Review*, 105(9), 2725–2756.

Amiti, M., and D. R. Davis (2012): "Trade, firms, and wages: Theory and evidence," *The Review of economic studies*, 79(1), 1–36.

Anderson, J., and E. van Wincoop (2004): "Trade Costs," *Journal of Economic Literature*, 42(3), 691–751.

Antràs, P., T. C. Fort, and F. Tintelnot (2014): "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms," Work in Progress.

Antràs, P., L. Garicano, and E. Rossi-Hansberg (2006): "Offshoring in a Knowledge Economy," *The Quarterly Journal of Economics*, 121(1), 31–77.

Arkolakis, C., and F. Esposito (2014): "Endogenous labor supply and the gains from international trade," *Processed, Yale University*.

Atkinson, A. B. (1970): "On the measurement of inequality," *Journal of Economic Theory*, 2(3), 244–263.

Attanasio, O., P. K. Goldberg, and N. Pavcnik (2004): "Trade Reforms and Wage Inequality in Colombia," *Journal of Development Economics*, 74, 331–66.

Attanasio, O., E. Hurst, and L. Pistaferri (2012): "The evolution of income, consumption, and leisure inequality in the US, 1980-2010," Discussion paper, National Bureau of Economic Research.

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103(6), 2121–68.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in U.S. Wage Inequality: Reassessing the Revisionists," *Review of Economics and Statistics*, 90(2), 300–23.

Benabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

Bergson, A. (1938): "A reformulation of certain aspects of welfare economics," *The Quarterly Journal of Economics*, pp. 310–334.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–90.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in International Trade," *Journal of Economic Perspectives*, 21(3), 105–130.

Broda, C., and J. Romalis (2008): "Inequality and Prices: Does China Benefit the Poor in America?," Chicago GSB mimeo.

Broda, C., and D. E. Weinstein (2006): "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 121(2), 541–85.

Burstein, A., and J. Vogel (2016): "International trade, technology, and the skill premium," *Journal of Political Economy*, forthcoming.

Carrère, C., A. Grujovic, and F. Robert-Nicoud (2015): "Trade and frictional unemployment in the global economy," .

Chang, B., Y. Chang, and S.-B. Kim (2016): "Pareto Weights in Practice: A Quantitative Analysis Across 32 OECD Countries," https://sites.google.com/site/yongsungchang/research.

Chetty, R. (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80(3), 969–1018.

Costinot, A., and A. Rodríguez-Clare (2013): "Trade Theory with Numbers: Quantifying the Consequences of Globalization," NBER Working Papers No. 18896.

Costinot, A., and J. Vogel (2010): "Matching and Inequality in the World Economy," *Journal of Political Economy*, 118(4), 747–786.

Das, S., M. J. Roberts, and J. R. Tybout (2007): "Market Entry Costs, Producer Heterogeneity, and Export Dynamics," *Econometrica*, 75(3), 837–73.

Davidson, C., and S. J. Matusz (2006): "Trade Liberalization And Compensation," *International Economic Review*, 47(3), 723–47.

Diamond, P., and E. Saez (2011): "The case for a progressive tax: from basic research to policy recommendations," *The Journal of Economic Perspectives*, 25(4), 165–190.

Diamond, P. A., and J. A. Mirrlees (1971): "Optimal Taxation and Public Production I: Production Efficiency," *The American Economic Review*, 61(1), 8–27.

Dixit, A. K., and V. Norman (1980): *Theory of International Trade*. Cambridge University Press.

——— (1986): "Gains from trade without lump-sum compensation," *Journal of International Economics*, 21(1-2), 111–22.

Eaton, J., S. Kortum, and F. Kramarz (2011): "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 79(5), 1453–1498.

Egger, H., and U. Kreickemeier (2009a): "Firm heterogeneity and the labor market effects of trade liberalization," *International Economic Review*, 50(1), 187–216.

——— (2009b): "Redistributing gains from globalisation," *The Scandinavian Journal of Economics*, 111(4), 765–788.

Ethier, W. J. (1982): "National and International Returns to Scale in the Modern Theory of International Trade," *The American Economic Review*, 72(3), 389–405.

Faggio, G., K. G. Salvanes, and J. Van Reenen (2007): "The Evolution of Inequality in Productivity and Wages: Panel Data Evidence," CEP Discussion Paper No. 821.

Fajgelbaum, P. D., and A. K. Khandelwal (2016): "Measuring the Unequal Gains from Trade*," *The Quarterly Journal of Economics*.

Feenstra, R. C., and G. H. Hanson (1999): "The Impact Of Outsourcing And High-Technology Capital On Wages: Estimates For The United States, 1979-1990," *Quarterly Journal of Economics*, 114(3), 907–40.

Feldstein, M. (1973): "On the optimal progressivity of the income tax," *Journal of Public Economics*, 2(4), 357–376.

Goldberg, P. K., and N. Pavcnik (2007): "Distributional Effects of Globalization in Developing Countries," *Journal of Economic Literature*, 45(1), 39–82.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 78(3), 402–417.

Guner, N., R. Kaygusuz, and G. Ventura (2014): "Income Taxation of U.S. Households: Facts and Parametric Estimates," *Review of Economic Dynamics*, 17(4), 559–581.

Harsanyi, J. C. (1953): "Cardinal Utility in Welfare Economics and in the Theory of Risk-taking," *Journal of Political Economy*, 61, 434–435.

Heathcote, J., K. Storesletten, and G. L. Violante (2014): "Optimal Tax Progressivity: An Analytical Framework," NBER Working Papers No. 19899.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2016): "Trade and Inequality: From Theory to Estimation," *Review of Economic Studies*, forthcoming.

Helpman, E., O. Itskhoki, and S. J. Redding (2010): "Inequality and Unemployment in a Global Economy," *Econometrica*, 78(4), 1239–1283.

Hendren, N. (2014): "The inequality deflator: Interpersonal comparisons without a social welfare function," Discussion paper, National Bureau of Economic Research.

Hicks, J. R. (1939): "The foundations of welfare economics," *The Economic Journal*, pp. 696–712.

Itskhoki, O. (2008): "Optimal Redistribution in an Open Economy," http://www.princeton.edu/~itskhoki/papers/RedistributionOpen.pdf.

Jakobsson, U. (1976): "On the measurement of the degree of progression," *Journal of Public Economics*, 5(1), 161–168.

Jaravel, X. (2016): "The Unequal Gains from Product Innovations," *mimeo Harvard University*.

Jones, C. I., and P. J. Klenow (2016): "Beyond GDP? Welfare across Countries and Time," *American Economic Review*, 106(9), 2426–57.

Kakwani, N. C. (1977): "Applications of Lorenz Curves in Economic Analysis," *Econometrica*, 45(3), 719–727.

Kaldor, N. (1939): "Welfare propositions of economics and interpersonal comparisons of utility," *The Economic Journal*, pp. 549–552.

Kaplow, L. (2004): "On the (Ir)Relevance of Distribution and Labor Supply Distortion to Government Policy," *The Journal of Economic Perspectives*, 18(4), 159–175.

Krueger, D., and F. Perri (2006): "Does income inequality lead to consumption inequality? Evidence and theory," *The Review of Economic Studies*, 73(1), 163–193.

Krugman, P. R. (1980): "Scale economies, product differentiation, and the pattern of trade," *American Economic Review*, 70(5), 950–9.

——— (2008): "Trade and Wages, Reconsidered," *Brookings Papers on Economic Activity*, 1(1), 103–37.

Lawrence, R. Z., and M. J. Slaughter (1993): "International Trade and American Wages in the 1980s: Giant Sucking Sound or Small Hiccup?," *Brookings Papers on Economic Activity*, 2, 161–226.

Layard, R., G. Mayraz, and S. Nickell (2008): "The marginal utility of income," *Journal of Public Economics*, 92(8), 1846–1857.

Levinsohn, J. (1999): "Employment Responses to International Liberalization in Chile," *Journal of International Economics*, 47, 321–44.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–725.

Melitz, M. J., and S. J. Redding (2015): "New Trade Models, New Welfare Implications," *American Economic Review*, 105(3), 1105–1046.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(2), 175–208.

Naito, H. (2006): "Redistribution, production inefficiency and decentralized efficiency," *International Tax and Public Finance*, 13(5), 625–640.

Piketty, T., and E. Saez (2003): "Income Inequality in the United States, 1913–1998," *The Quarterly journal of economics*, 118(1), 1–41.

Porto, G. (2015): "Trading off Poverty, Inequality and Growth with Trade Policy," Discussion paper, Preliminary Slides, Universidad Nacional de La Plata.

ROBBINS, L. (1932): "The Nature and Significance of Economic Science," *The Philosophy of Economics*, pp. 113–40.

RODRIGUEZ-CLARE, A., S. GALLE, AND M. YI (2015): "Slicing the Pie: Quantifying the Aggregate and Distributional Consequences of Trade," Discussion paper, UC Berkeley, Preliminary Draft.

RODRIK, D. (1992): "The rush to free trade in the developing world: Why so late? Why now? Will it last?," Discussion paper, National Bureau of Economic Research.

ROTHSCHILD, M., AND J. E. STIGLITZ (1970): "Increasing risk: I. A definition," *Journal of Economic theory*, 2(3), 225–243.

SACHS, D., A. TSYVINSKI, AND N. WERQUIN (2016): "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium," .

SAEZ, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–29.

——— (2002): "Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117(3), 1039–73.

SAMPSON, T. (2014): "Selection into trade and wage inequality," *American Economic Journal: Microeconomics*, 6(3), 157–202.

SAMUELSON, P. A. (1948): "Foundations of economic analysis," .

SPECTOR, D. (2001): "Is it possible to redistribute the gains from trade using income taxation?," *Journal of International Economics*, 55.

STIGLITZ, J. E. (1982): "Self-selection and Pareto efficient taxation," *Journal of Public Economics*, 17(2), 213–40.

VICKREY, W. (1945): "Measuring Marginal Utility by Reactions to Risk," *Econometrica*, 13(4), 319–333.

YEAPLE, S. R. (2005): "A simple model of firm heterogeneity, international trade, and wages," *Journal of International Economics*, 65(1), 1–20.

A Appendix

A.1 Properties of the Welfarist and Costly-Redistribution Corrections

In this Appendix, we discuss certain properties of the welfarist and costly-redistribution corrections that hold for general distributions of income.

Let us begin with the welfarist correction Δ in equation (8). The fact that that $\Delta \leq 1$ follows immediately from Jensen's inequality, while it is obvious that $\Delta = 1$ only if either there is no inequality aversion ($\rho = 0$) or if the distribution of disposable income has zero dispersion (so $\mathbb{E}\left(r_\varphi^d\right)^{1-\rho} = \left(\mathbb{E}r_\varphi^d\right)^{1-\rho}$). The fact that Δ is reduced by mean-preserving spreads of the distribution of disposable income was proven by Atkinson (1970) invoking the results in Rothschild and Stiglitz (1970). That Δ is decreasing in ρ, holding constant the distribution of disposable income, can be established using Jensen's inequality:

$$\Delta(F,\rho') = \frac{\left[\mathbb{E}\left(r_\varphi^d\right)^{1-\rho'}\right]^{\frac{1}{1-\rho'}}}{\mathbb{E}r_\varphi^d} = \Delta(F;\rho) \times \left(\frac{\left[\mathbb{E}x^{\upsilon}\right]^{1/\upsilon}}{\mathbb{E}x}\right)^{\frac{1}{1-\rho}} < \Delta(F;\rho),$$

where $x \equiv \left(r_\varphi^d\right)^{1-\rho}$ and $\upsilon \equiv (1-\rho')/(1-\rho) \in (0,1)$ for $\rho' > \rho$.

The following result is useful for determining the signs of comparative statics with respect to ϕ.

Lemma 1. *Let x be a random variable and $g(\phi) : [0,1] \to \Re_+$ such that $g(\phi) > 0$ in $[0,1)$ and $g'(\phi) \leq 0$. For any positive real number $\upsilon > 0$, we have*

$$\frac{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\ln(x)\right]}{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\right]} < \frac{\mathbb{E}_x\left[x^{g(\phi)}\ln(x)\right]}{\mathbb{E}_x\left[x^{g(\phi)}\right]}$$

Proof. *The two ratios of expectations coincide at $\upsilon = 0$ and the left-hand side falls with increases in υ since*

$$\frac{\partial}{\partial\upsilon}\frac{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\ln(x)\right]}{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\right]} = -g\left(\phi\right)\left\{\frac{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\left(\ln(x)\right)^2\right]}{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\right]} - \left(\frac{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\ln(x)\right]}{\mathbb{E}_x\left[x^{g(\phi)(1-\upsilon)}\right]}\right)^2\right\} < 0$$

where the strict inequality follows from Jensen's inequality, since $h(z) = z^2$ is a strictly convex function.

In our closed-economy model in section 3.2, we claimed that the welfarist correction Δ is increasing in ϕ. To see this note from equation (19) that r_φ^d is proportional to $\varphi^{\frac{\beta(1+\varepsilon)(1-\phi)}{1+\varepsilon\phi}}$ so that the derivative with respect to ϕ equals

$$\frac{\partial\Delta}{\partial\phi} = -\Delta\left\{\frac{\mathbb{E}_\varphi\left[^{g(\phi)(1-\rho)}\ln(\varphi)\right]}{\mathbb{E}_\varphi\left[\varphi^{g(\phi)(1-\rho)}\right]} - \frac{\mathbb{E}_\varphi\left[\varphi^{g(\phi)}\ln(\varphi)\right]}{\mathbb{E}_\varphi\left[\varphi^{g(\phi)}\right]}\right\}$$

with $g(\phi) = \frac{\beta(1+\varepsilon)(1-\phi)}{1+\varepsilon\phi}$. That this derivative is positive is then immediate from applying Lemma 1 for the case $\upsilon = \rho$.

Finally, and still with regards to the welfarist correction, we claimed at the end of section 2.3 that if one were to compute the percentage change in all agents' consumption $c_\varphi = r_\varphi^d$ that would make society indifferent between F_r^d and $F_r^{d\prime}$, the answer one would get would be $\mu^C = (1+\mu^R) \times \frac{\Delta'}{\Delta} - 1$ regardless of whether social welfare is measured in terms of the function V or of any monotonic transformation of

V (including W). To prove this, note that for any nondecreasing function f,

$$f\left(\int \frac{\left(\left(1+\mu^C\right)c_\varphi\right)^{1-\rho}-1}{1-\rho}\mathrm{d}H_\varphi\right) = f\left(\int \frac{\left(c'_\varphi\right)^{1-\rho}-1}{1-\rho}\mathrm{d}H_\varphi\right)$$

only if $1+\mu^C = \left[\mathbb{E}\left(c'_\varphi\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} / \left[\mathbb{E}\left(c_\varphi\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} = \Delta' R' / \left(\Delta R\right) = (1+\mu^R)\Delta'/\Delta$.

Let us next turn to the costly-redistribution correction Θ in equation (14). It is immediate that Hölder's inequality implies that $\Theta \leq 1$, and that $\Theta = 1$ if and only if the tax-transfer system features zero progressivity ($\phi = 0$) or if the elasticity of taxable income is zero ($\varepsilon = 0$). Less trivially, we can also show that Θ is strictly decreasing in the tax progressivity rate ϕ for any primitive distribution of potential income $\tilde{r}_\varphi$, which is the income agent φ would obtain in a counterfactual economy without progressive redistribution (i.e., with $\phi = 0$ in the tax schedule). To see this, first note that the progressivity rule (11) and the constant elasticity assumption (12) imply that $r_\varphi \propto (\tilde{r}_\varphi)^{\frac{1}{1+\varepsilon\phi}}$. Plugging this into equation (14) we obtain that

$$\frac{\partial \ln \Theta}{\partial \phi} = -\frac{\varepsilon}{1-\phi} - (1+\varepsilon)\frac{\partial}{\partial \phi}\left(\ln \mathbb{E}\left[\tilde{r}_\varphi^{\frac{1-\phi}{1+\varepsilon\phi}}\right]^{\frac{\varepsilon}{1+\varepsilon}} - \ln \mathbb{E}\left[\tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\right]\right)$$

The first term is clearly negative while the second term is negative if and only if

$$\frac{\mathbb{E}\left[\tilde{r}_\varphi^{\frac{1-\phi}{1+\varepsilon\phi}}\ln(\tilde{r}_\varphi)\right]}{\mathbb{E}\left[\tilde{r}_\varphi^{\frac{1-\phi}{1+\varepsilon\phi}}\right]} < \frac{\mathbb{E}\left[\tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\ln(\tilde{r}_\varphi)\right]}{\mathbb{E}\left[\tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\right]}$$

We can then again apply Lemma 1 with $g(\phi) = 1/(1+\varepsilon\phi)$ and $\upsilon = \phi$ to verify that this inequality indeed holds for any $\phi > 0$. As a result, both terms in $\frac{\partial \ln \Theta}{\partial \phi}$ are negative and Θ is thus decreasing in ϕ.

Consider next the costly-redistribution correction in the constant elasticity model in which ability is fixed and income is endogenous. Remember that in that case, Θ is a function of income, which is now endogenous, and it also appears to the power κ and is pre-multiplied by the term $(1+\varepsilon\phi)$ in (24). Note, however, that from equation (19), pre-tax income r_φ is proportional to $\varphi^{\frac{\beta(1+\varepsilon)}{1+\varepsilon\phi}}$, setting $\phi = 0$ we obtain that potential income $\tilde{r}_\varphi$ is proportional to $\varphi^{\beta(1+\varepsilon)}$. Hence, and quite intuitively, ϕ plays no role in the mapping between ability and potential income (at least up to a constant that would cancel in the ratio of expectations in equation (14)). We can now invoke our results above and state that Θ necessarily continues to decline in ϕ (despite the endogeneity of income). We can thus focus on the term $(1+\varepsilon\phi)(1-\phi)^{\varepsilon\kappa}$, for which:

$$\frac{\partial(1+\varepsilon\phi)(1-\phi)^{\varepsilon\kappa}}{\partial \phi} = \varepsilon(1-\phi)^{\varepsilon\kappa}\left(1-\kappa\frac{1+\varepsilon\phi}{1-\phi}\right) < 0,$$

where the negative sign follows from $\kappa \geq 1$. This establishes that $(1+\varepsilon\phi)\Theta^\kappa$ is decreasing in ϕ.

A corollary of this result, our previous result that Δ increases in ϕ (as ϕ reduces the dispersion of the after-tax incomes), and the welfare decomposition in (24) is that social welfare is shaped by the product of two terms affected by ϕ in opposite directions. As argued in the main text, when setting the optimal degree of tax progressivity ϕ, a social planner would thus seek to balance these two conflicting forces.

We conclude this Appendix by formally showing that, when considering two distributions of income F_r and F'_r, we have that $\Theta(F'_r, \phi, \varepsilon) < \Theta(F_r, \phi, \varepsilon)$ when F'_r is a mean preserving multiplicative spread of F_r. Remember that the distribution F'_r is a mean preserving multiplicative spread of F_r whenever there exists a random variable θ independent of the original income r such that $r' = (1+\theta)\,r$ with $\mathbb{E}\,(\theta) = 0$.

We thus can write:

$$\begin{aligned}
\Theta(F_r',\phi,\varepsilon) &= (1-\phi)^{\varepsilon}\frac{\left(\mathbb{E}r_\varphi'\right)^{1+\varepsilon}}{\left(\mathbb{E}\left(r_\varphi'\right)^{1-\phi}\right)^{\varepsilon}\cdot\left(\mathbb{E}\left(r_\varphi'\right)^{1+\varepsilon\phi}\right)} = \\
&= (1-\phi)^{\varepsilon}\frac{\left(\mathbb{E}r_\varphi\right)^{1+\varepsilon}}{\left(\mathbb{E}\left(r_\varphi\right)^{1-\phi}\right)^{\varepsilon}\cdot\left(\mathbb{E}\left(r_\varphi\right)^{1+\varepsilon\phi}\right)}\frac{\left(\mathbb{E}\left(1+\theta\right)\right)^{1+\varepsilon}}{\left(\mathbb{E}\left(1+\theta\right)^{1-\phi}\right)^{\varepsilon}\cdot\left(\mathbb{E}\left(1+\theta\right)^{1+\varepsilon\phi}\right)} \\
&\leq (1-\phi)^{\varepsilon}\frac{\left(\mathbb{E}r_\varphi\right)^{1+\varepsilon}}{\left(\mathbb{E}\left(r_\varphi\right)^{1-\phi}\right)^{\varepsilon}\cdot\left(\mathbb{E}\left(r_\varphi\right)^{1+\varepsilon\phi}\right)} = \Theta(F_r,\phi,\varepsilon)
\end{aligned}$$

where $\left(\mathbb{E}\left(1+\theta\right)\right)^{1+\varepsilon} \leq \left(\mathbb{E}\left(1+\theta\right)^{1-\phi}\right)^{\varepsilon}\cdot\left(\mathbb{E}\left(1+\theta\right)^{1+\varepsilon\phi}\right)$ follows from Hölder's inequality.

A.2 Two Parametric Examples: Lognormal and Pareto

In this Appendix, we consider two common parametric examples to further illustrate the properties of the correction terms introduced in section 2. Specifically, we consider the cases in which the distribution of market income is either lognormal or Pareto. Even though neither of these two distributions matches observed incomes perfectly, these are the two most popular distributions in the literature offering a reasonably good fit of the data.[38] In both cases, we postulate a distribution for (before-tax) market incomes r, and calculate the (after-tax) disposable income according to (11) for a given value of ϕ, that is $r^d = kr^{1-\phi}$.

Lognormal distribution When market incomes are distributed lognormally with a mean parameter μ and a variance parameter σ^2, the after-tax disposable income is also distributed log-normally with variance parameter $(1-\phi)^2\sigma^2$. In this case, it is straightforward to show that the welfarist and costly-redistribution corrections are equal to:

$$\Delta = \Delta\left(\sigma;\rho,\phi\right) = \exp\left\{-\rho(1-\phi)^2\frac{\sigma^2}{2}\right\}, \tag{37}$$

$$\Theta = \Theta(\sigma;\varepsilon,\phi) = (1-\phi)^{\varepsilon}\exp\left\{-\varepsilon(1+\varepsilon)\phi^2\frac{\sigma^2}{2}\right\}. \tag{38}$$

Thus, in both cases, the size of the corrections is increasing in the single parameter σ^2 governing the inequality of income. Furthermore, the effect of inequality on the welfarist correction is magnified by a higher inequality aversion ρ and moderated by the extent of tax progressivity ϕ. In contrast, the effect of inequality on the costly redistribution correction is magnified by a higher degree of progressivity ϕ, and also by a higher taxable income elasticity ε. Note also that because the Gini coefficient associated with a lognormal distribution is simply given by $G = 2\Phi\left(\sigma/\sqrt{2}\right) - 1$, it is straightforward to re-express (37) and (38) as functions of the Gini coefficient rather than σ^2.

Pareto distribution When market incomes are distributed Pareto with shape parameter α (with $\alpha > 1$ ensuring the existence of the first moments and $\alpha > 2$ the existence of the second moments), the

[38] It is often argued that the Pareto distribution provides a good fit of the top percentiles of the income distribution, while the bottom 80–90% of incomes are better approximated by a lognormal distribution. It is straightforward to develop analogous formulas for mixtures of lognormal and Pareto distributions, such as the case of the double Pareto-lognormal distribution, which offers a better fit of the data.

after-tax disposable income is distributed Pareto with shape parameter $\alpha/(1-\phi)$, with a lower value of α corresponding to greater inequality.[39] In this case, we obtain:

$$\Delta = \Delta(\alpha;\rho,\phi) = \frac{\alpha-(1-\phi)}{\alpha}\left[\frac{\alpha}{\alpha-(1-\rho)(1-\phi)}\right]^{\frac{1}{1-\rho}}, \tag{39}$$

$$\Theta = \Theta(\alpha;\varepsilon,\phi) = (1-\phi)^{\varepsilon}\left[1+\frac{\phi}{\alpha-1}\right]^{\varepsilon}\left[1-\frac{\varepsilon\phi}{\alpha-1}\right]. \tag{40}$$

Straightforward differentiation demonstrates that both Δ and Θ are increasing in the shape parameter α, and thus higher inequality levels (smaller α) are associated with larger inequality corrections (smaller Δ and Θ). Because the Gini coefficient of a Pareto distribution is simply given by $G = 1/(2\alpha-1)$, it is again trivial to re-express (38) and (40) as functions of the Gini coefficient rather than α. Furthermore, it can also be easily verified that Δ is again decreasing in inequality aversion ρ and increasing in the degree of tax progressivity ϕ, while Θ is instead decreasing in ϕ and also decreases in the taxable income elasticity ε.

These two parametric examples illustrate that, on account of both the welfarist and costly-redistribution corrections, social welfare is negatively impacted by higher (or increasing) levels of inequality of the income distribution. Nonetheless, the two measures behave differently with regards to a key tool available to governments to correct such inequality, namely progressive taxation.

A.3 Details on the Derivations in Section 3

Start with equation (19):

$$r_{\varphi} = \left[\beta(1-\phi)k\right]^{\frac{\varepsilon}{1+\varepsilon\phi}}\left[Q^{1-\beta}\varphi^{\beta}\right]^{\frac{1+\varepsilon}{1+\varepsilon\phi}}$$

which implies the potential income (i.e., income in the absence of tax progressivity, that is r_{φ} evaluated at $\phi = 0$) given by:

$$\tilde{r}_{\varphi} = \left[\beta\tilde{k}\right]^{\varepsilon}\left[\tilde{Q}^{1-\beta}\varphi^{\beta}\right]^{1+\varepsilon}.$$

We first solve for the general equilibrium variables under zero taxes, $(\tilde{k},\tilde{Q})$. We have from (18) that

$$\tilde{k} = (1-g),$$

since $Q = R = \int r_{\varphi}\mathrm{d}H_{\varphi}$, which further implies:

$$\tilde{Q} = \int \left[\beta\tilde{k}\right]^{\varepsilon}\left[\tilde{Q}^{1-\beta}\varphi^{\beta}\right]^{1+\varepsilon}\mathrm{d}H_{\varphi}.$$

Solving for $\tilde{Q}$ we get:

$$\tilde{Q} = \left[\beta(1-g)\right]^{\kappa\varepsilon}\left(\int \varphi^{\beta(1+\varepsilon)}\mathrm{d}H_{\varphi}\right)^{\kappa}, \qquad \text{where} \qquad \kappa \equiv \frac{1}{1-(1-\beta)(1+\varepsilon)}.$$

Therefore we can write the solution for $\tilde{r}_{\varphi}$ without endogenous variables as:

$$\tilde{r}_{\varphi} = \left[\beta(1-g)\right]^{\kappa\varepsilon}\left(\int \varphi^{\beta(1+\varepsilon)}\mathrm{d}H_{\varphi}\right)^{\kappa}\frac{\varphi^{\beta(1+\varepsilon)}}{\int \varphi^{\beta(1+\varepsilon)}\mathrm{d}H_{\varphi}} = \tilde{Q}\frac{\varphi^{\beta(1+\varepsilon)}}{\int \varphi^{\beta(1+\varepsilon)}\mathrm{d}H_{\varphi}}.$$

[39]Specifically, the cdf of the pre-tax market income in this case is $F_r = 1-(r_{\min}/r)^{\alpha}$.

Note that an increase in g decreases revenues for all agents with an elasticity $\kappa\varepsilon > \varepsilon$ due to the CES (love of variety) demand externality when $\beta < 1$.

We use the above derivations as interim steps to characterizing the allocation for $\phi > 0$. Note that we can write:

$$r_\varphi = (1-\phi)^{\frac{\varepsilon}{1+\varepsilon\phi}} \left(\frac{k}{1-g}\right)^{\frac{\varepsilon}{1+\varepsilon\phi}} \left(\frac{Q}{\tilde{Q}}\right)^{\frac{(1-\beta)(1+\varepsilon)}{1+\varepsilon\phi}} \tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}},$$

$$\frac{k}{1-g} = \frac{\int r_\varphi \mathrm{d}H_\varphi}{\int r_\varphi^{1-\phi}\mathrm{d}H_\varphi} = (1-\phi)^{\frac{\varepsilon\phi}{1+\varepsilon\phi}} \left(\frac{k}{1-g}\right)^{\frac{\varepsilon\phi}{1+\varepsilon\phi}} \left(\frac{Q}{\tilde{Q}}\right)^{\frac{(1-\beta)(1+\varepsilon)\phi}{1+\varepsilon\phi}} \frac{\int \tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\mathrm{d}H_\varphi}{\int \tilde{r}_\varphi^{\frac{1-\phi}{1+\varepsilon\phi}}\mathrm{d}H_\varphi},$$

$$Q = (1-\phi)^{\frac{\varepsilon}{1+\varepsilon\phi}} \left(\frac{k}{1-g}\right)^{\frac{\varepsilon}{1+\varepsilon\phi}} \left(\frac{Q}{\tilde{Q}}\right)^{\frac{(1-\beta)(1+\varepsilon)}{1+\varepsilon\phi}} \int \tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\mathrm{d}H_\varphi.$$

Solving out $k/(1-g)$, we obtain:

$$\frac{k}{1-g} = (1-\phi)^{\varepsilon\phi}\left(\frac{Q}{\tilde{Q}}\right)^{(1-\beta)(1+\varepsilon)\phi}\left(\frac{\int \tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\mathrm{d}H_\varphi}{\int \tilde{r}_\varphi^{\frac{1-\phi}{1+\varepsilon\phi}}\mathrm{d}H_\varphi}\right)^{1+\varepsilon\phi}$$

and substituting this into the expression for $Q = R$:

$$\frac{R}{\tilde{R}} = \frac{Q}{\tilde{Q}} = (1-\phi)^{\kappa\varepsilon}\left[\frac{\left(\int \tilde{r}_\varphi^{\frac{1}{1+\varepsilon\phi}}\mathrm{d}H_\varphi\right)^{1+\varepsilon}}{\left(\int \tilde{r}_\varphi^{\frac{1-\phi}{1+\varepsilon\phi}}\mathrm{d}H_\varphi\right)^{\varepsilon}\left(\int \tilde{r}_\varphi \mathrm{d}H_\varphi\right)}\right]^{\kappa} = \Theta^\kappa.$$

A.4 Proofs of Main Theoretical Results

Proof of Proposition 1

Beginning with equation (20), set $\phi = 0$ and $k = 1-g$, so $\tilde{u}_\varphi = \frac{1-g}{1+\varepsilon}\tilde{r}_\varphi$. Aggregating over individuals using (22) with $\rho = 0$, and noting $\tilde{Q} = \tilde{R} = \int \tilde{r}_\varphi dH_\varphi$, results in equation (23) in the Proposition.

Proof of Proposition 2

Note first that we can write (22) as

$$W = \frac{\left(\int u_\varphi^{1-\rho}\mathrm{d}H_\varphi\right)^{1/(1-\rho)}}{\int u_\varphi \mathrm{d}H_\varphi} \times \int u_\varphi \mathrm{d}H_\varphi.$$

Plugging (20), and invoking (11) and (18), we can simplify this to

$$W = \frac{1+\varepsilon\phi}{1+\varepsilon}\frac{\left(\int \left(r_\varphi^d\right)^{1-\rho}\mathrm{d}H_\varphi\right)^{1/(1-\rho)}}{\int r_\varphi^d \mathrm{d}H_\varphi}(1-g)Q.$$

Invoking (21), as well as the definition of Δ in (9), we thus have:

$$W = \Delta \times (1-g)\frac{1+\varepsilon\phi}{1+\varepsilon}\Theta^\kappa \times \tilde{Q}.$$

Equation (24) is then obtained by plugging equation (23) in Proposition 1.

Proof of Proposition 3

Take two individuals with ability $\varphi^H > \varphi^L$. From equations (30) and (32), we have that pre-tax incomes satisfy:

$$\frac{r_{\varphi^H}}{r_{\varphi^L}} = \left(\frac{\Upsilon_{n_{\varphi^H}}}{\Upsilon_{n_{\varphi^L}}}\right)^{\frac{(1+\varepsilon)(1-\beta)}{1+\varepsilon\phi}} \left(\frac{\varphi^H}{\varphi^L}\right)^{\frac{(1+\varepsilon)\beta}{1+\varepsilon\phi}}. \tag{41}$$

The second term is identical as in the closed-economy model, while the first term is new, and because n_φ is nondecreasing in φ, this term is necessarily (weakly) higher than 1, and it will be larger than one as long as some individuals export in some but not all markets. The proof for after-tax incomes is therefore immediate since $r_\varphi - T\big(r_\varphi\big) = kr_\varphi^{1-\phi}$.

To show the result for the case of utility levels, we begin by using (34) to define an ability level $\bar{\varphi}_n$ such that

$$\left[\Upsilon_n^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \Upsilon_{n-1}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}}\right] u_{0\bar{\varphi}_n} = Fn^\alpha.$$

In words, $\bar{\varphi}_n$ is the minimum ability level such that choosing to export in n foreign markets is optimal. Note then that equation (31) can be expressed as

$$u_\varphi = \Psi\left(\varphi\right) u_{0\varphi}$$

where $u_{0\varphi}$ is defined in (33) and

$$\Psi\left(\varphi\right) = \left(\Upsilon_{n_\varphi}\right)^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \sum_{n=1}^{n_\varphi}\left[\Upsilon_n^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \Upsilon_{n-1}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}}\right]\left(\frac{\bar{\varphi}_n}{\varphi}\right)^{\frac{(1+\varepsilon)\beta(1-\phi)}{1+\varepsilon\phi}}. \tag{42}$$

We next show that $\Psi\left(\varphi\right) \geq 1$ and $\Psi'\left(\varphi\right) \geq 0$ which guarantees the validity of the statement in Proposition 3 with regards to utility levels. Note that $\Psi\left(\varphi\right) = 1$ for the lowest ability levels for which $n_\varphi = 0$, so it suffices to show that $\Psi'\left(\varphi\right) \geq 0$. This is obvious in the interval of abilities for which a common n_φ is optimal. In other words, for any $\varphi^H > \varphi^L$ for which $n_{\varphi^H} = n_{\varphi^L}$. Whenever $n_{\varphi^H} = n_{\varphi^L} + 1$, notice that

$$\begin{aligned}\Psi\left(\varphi^H\right) - \Psi\left(\varphi^L\right) = &\left[\Upsilon_{n_\varphi}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \Upsilon_{n_\varphi-1}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}}\right]\left[1 - \left(\frac{\bar{\varphi}_{n_\varphi}}{\varphi^H}\right)^{\frac{(1+\varepsilon)\beta(1-\phi)}{1+\varepsilon\phi}}\right] \\ &+ \sum_{n=1}^{n_\varphi-1}\left[\Upsilon_n^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}} - \Upsilon_{n-1}^{\frac{(1+\varepsilon)(1-\beta)(1-\phi)}{1+\varepsilon\phi}}\right]\left[\left(\frac{\bar{\varphi}_n}{\varphi^L}\right)^{\frac{(1+\varepsilon)\beta(1-\phi)}{1+\varepsilon\phi}} - \left(\frac{\bar{\varphi}_n}{\varphi^H}\right)^{\frac{(1+\varepsilon)\beta(1-\phi)}{1+\varepsilon\phi}}\right] > 0.\end{aligned}$$

It is then straightforward to show that the same is true for $n_{\varphi^H} = n_{\varphi^L} + m$ for any $m > 1$. This implies that $\Psi\left(\varphi\right)$ is nondecreasing and strictly higher than one as long as some individuals export in some but not all markets.

Proof of Proposition 4

By Proposition 3, we can focus on showing that the real income and utility of the lowest-ability individuals who only sell locally increases in moving from autarky to a trade equilibrium. Appealing to equations (32) and (33), it then suffices to show that both k and Q are raised by trade. To show that both k and Q increase with trade, begin with the individual-level optimization problem of setting $\ell_\varphi \in \mathbb{R}^+$ and

$n_\varphi \in \{0, 1, ..., N\}$ to maximize

$$u_\varphi = k\left(Q^{1-\beta}\Upsilon_{n_\varphi}^{1-\beta}(\varphi\ell_\varphi)^\beta\right)^{1-\phi} - \frac{1}{\gamma}\ell_\varphi^\gamma - f_x\sum_{n=1}^{n_\varphi} n^\alpha$$

with $\Upsilon_{n_\varphi} = 1 + n_\varphi\tau^{-\frac{\beta}{1-\beta}}$ and $n_\varphi \in \{0, 1, ..., N\}$. By standard monotone comparative statics arguments, holding constant k and Q, the choices of ℓ_φ and n_φ are both higher in a trade equilibrium in which τ is bounded than in an autarky equilibrium in which $\tau = +\infty$. These increases will in turn increase aggregate income $Q = \int r_\varphi(Q,k)\,\mathrm{d}H_\varphi$, and holding constant k, further increase the choices of ℓ_φ and n_φ. This is the love-for-variety magnification effect highlighted in the closed-economy version of the model, and stability requires that $(1-\beta)(1+\varepsilon) < 1$. Turning to k, notice that we can write

$$k = (1-g)\frac{\int r_\varphi(Q,k)\,\mathrm{d}H_\varphi}{\int (r_\varphi(Q,k))^{1-\phi}\,\mathrm{d}H_\varphi} = (1-g)\,Q^\phi\frac{\left(\mathbb{E}r_\varphi\right)^{1-\phi}}{\left(\mathbb{E}r_\varphi^{1-\phi}\right)}.$$

That k increases from trade then follows from Q increasing and from the fact that the increase in inequality formalized in Proposition 3 will also increase the last term in k. To see this more formally, simply note that Proposition 3 together with the increase in r_φ for the lowest-ability individual (holding constant k) implies that trade leads to an income distribution that first-order (and thus second-order) stochastically dominates the autarkic one, and we can thus invoke the results in Atkinson (1970) to argue that $\left(\mathbb{E}r_\varphi\right)^{1-\phi}/\left(\mathbb{E}r_\varphi^{1-\phi}\right)$ is larger with trade.

Proof of Proposition 5

We need to show that

$$\Delta = \frac{\left[\mathbb{E}\left(u_\varphi\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}}{\mathbb{E}u_\varphi}$$

is lower under a trade equilibrium in which some (but not all) individuals export to some markets than under autarky, and that Δ is strictly decreasing in ρ. The proof is a direct corollary of Propositions 3 and 4, and the results in Atkinson (1970). More specifically, Propositions 3 and 4 imply that the distribution of utility in a trade equilibrium in which some (but not all) individuals export to some markets first-order (and thus second-order) stochastically dominates that under autarky. It then follows from the results in Atkinson (1970) that Δ is lower in a trade equilibrium, and more so the higher is ρ.

A.5 Details of the Numerical Analysis

The NBER-IRS data reports the pre-tax income distribution. Define an observation as r_i and let M be the total number of observations. We first show how to solve for the equilibrium of the trade model and how to calibrate the ability distribution φ such that the model delivers a pre-tax income distribution $r_\varphi \overset{d}{=} r_i$, conditional on a set of parameters $\{\gamma, \beta, \phi, N, \tau, f_x, \alpha\}$. We then discuss the moments used to calibrate the trade parameters $\{\tau, f_x, \alpha\}$.

A.5.1 Ability Calibration and Equilibrium Computation

The general equilibrium variables Q and k can be read off the data as

$$Q = \frac{1}{M}\sum_{i=1}^{M} r_i$$
$$k = \sum_{i=1}^{M} r_i / \sum_{i=1}^{M} r_i^{1-\phi}$$

Define $\bar{\varphi}_n$ as the threshold ability such that an agent is indifferent between exporting to n or $n-1$ locations. In terms of utility the following condition holds

$$\begin{aligned} & k\left(\Upsilon_n^{1-\beta} Q^{1-\beta}\left(\bar{\varphi}_n \ell_n(\bar{\varphi}_n)\right)^{\beta}\right)^{1-\phi} - \sum_{m=1}^{n} f_x m^{\alpha} - \frac{1}{\gamma}\ell_n(\bar{\varphi}_n)^{\gamma} \\ & = k\left(\Upsilon_{n-1}^{1-\beta} Q^{1-\beta}\left(\bar{\varphi}_n \ell_{n-1}(\bar{\varphi}_n)\right)^{\beta}\right)^{1-\phi} - \sum_{m=1}^{n-1} f_x m^{\alpha} - \frac{1}{\gamma}\ell_{n-1}(\bar{\varphi}_n)^{\gamma} \end{aligned} \tag{43}$$

where $\ell_n(\varphi)$ is the optimal labor allocation of an agent with ability φ and exporting to n locations. From the first order condition

$$\ell_m(\varphi) = \left[\beta\left(1-\phi\right) k \left(\Upsilon_m^{1-\beta} Q^{1-\beta} \varphi^{\beta}\right)^{1-\phi}\right]^{\frac{1}{\gamma-\beta(1-\phi)}}$$

So that equation (43) pins down $\bar{\varphi}_n$ at each $n = 1, \ldots, N$.

Reduce the continuum of agents to a finite number M and index them by i. In equilibrium, the labor allocation of agent i equals

$$\ell_i = \left[\beta\left(1-\phi\right) k r_i^{1-\phi}\right]^{\frac{1}{\gamma}}$$

Using the definition of revenue (26), ability is given for any n as

$$\varphi_i(n) = \frac{1}{\ell_i}\left(\frac{r_i}{\Upsilon_n^{1-\beta} Q^{1-\beta}}\right)^{\frac{1}{\beta}}$$

The ability level of i is defined as $\varphi_i = \varphi_i(n_i)$ with $n_i = n$ such that $\varphi_i(n) \in (\bar{\varphi}_n, \bar{\varphi}_{n+1})$. In other words, for every agent $\varphi_i(n)$ is the ability level that is consistent with exporting to an arbitrary n locations and delivering revenue r_i. The number of locations to which an agent actually exports is only consistent with the ability level draw if it is within the threshold in which the agent would freely choose that number of export locations.

A.5.2 Trade Parameters Calibration

The previous subsection has shown how to obtain an equilibrium which delivers endogenously the pre-tax income distribution r_i for any set of parameters $\{\gamma, \beta, \phi, N, \tau, f_x, \alpha\}$. We now define the moments used to determine the values of the trade parameters conditional on $\{\gamma, \beta, \phi, N\}$. Define the simulated moments

$$M1 = \frac{\sum_i \left(\frac{\Upsilon_{n_i}-1}{\Upsilon_{n_i}}\right) r_i}{\sum_i r_i}$$
$$M2 = \frac{\sum_{i:n_i>0} r_i}{\sum_i r_i}$$
$$M3 = \frac{\sum_{i:n_i>1} \left(\frac{\Upsilon_{n_i}-1}{\Upsilon_{n_i}}\right) r_i}{\sum_{i:n_i>0} \left(\frac{\Upsilon_{n_i}-1}{\Upsilon_{n_i}}\right) r_i}$$

The first moment corresponds to the trade share and is aggregate exports over aggregate output. The data analog comes from U.S. National Income and Product Accounts which report gross trade and gross output for the U.S. economy. The ratio of these delivers a trade share of 7.7% for 2007. The second moment is aggregate output of exporters as a share of aggregate output. We target a value of 61.8% which corresponds to the ratio of sales of exporters to sales of all firms in U.S. Census data, see Antràs, Fort, and Tintelnot, 2014. Finally, the last moment is the share of aggregate exports that are produced by households exporting to more than one location. The U.S. Census Bureau reports that 88.9% of exports in 2007 were accounted for by firms exporting to at least 5 destinations. Since $N = 5$ so that there are 6 destinations in our numerical simulations with the calibration corresponding to the U.S. economy we interpret an agent exporting to more than one destination as roughly equivalent to exporting to five or more countries in the data. Hence we use the share 88.9% for the third moment.

These three moments pin down the 2007 trade equilibrium parameters τ, f_x, and α, respectively, by equating the simulated moments to the values discussed above. The iceberg trade cost for 1979 is calibrated utilizing the pre-tax income distribution for 1979 together with the other parameters but targeting a trade share of 4.9%.