

Inequality as an Externality: Existence and Consequences

Morten Nyborg Stostad

▶ **To cite this version:**

Morten Nyborg Stostad. Inequality as an Externality: Existence and Consequences. Economics and Finance. 2019. dumas-02407577

HAL Id: dumas-02407577

https://dumas.ccsd.cnrs.fr/dumas-02407577

Submitted on 12 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

MASTER THESIS N° 2019 – 05

Inequality as an Externality: Existence and Consequences

Morten Nyborg Stostad

JEL Codes: H21, H23, D62, D63
Keywords:

Inequality as an Externality: Existence and Consequences

Morten Nyborg Støstad
Supervisor: Daniel Waldenström

Paris School of Economics
Analysis and Policy in Economics
Master's Thesis

June 2, 2019

Abstract

Does inequality shape society? This thesis proposes to consider inequality as an economic externality by inserting an inequality measure into the individual utility function. The externality is largely distinct from other conventional inequality considerations, and directly captures both (i) societal effects of inequality and (ii) individual distributional preferences. It has large theoretical implications: the Atkinson-Stiglitz theorem no longer holds, as goods consumed by low- and high-income agents are optimally subsidized and taxed, respectively. The First Welfare theorem falls away, and the Mirrlees (1971) optimal income taxation model changes drastically. An appropriate externality can benefit public policy by reconciling differences between economic theory and current tax systems. To motivate the discussion, the existing empirical evidence for an inequality externality is critically examined.

JEL Codes: H21, H23, D62, D63



1 Introduction

The academic research on inequality is rapidly progressing. Resources are increasingly devoted to the measurement of inequality levels, and the causes and empirical consequences of inequality are being examined more than ever.

Similarly, the global focus on inequality is growing. Inequality itself is increasingly seen as a public bad – or sometimes as a public good. There is a growing belief that the level of inequality is a defining characteristic of a society, potentially affecting anything from public health to innovation to political polarization. The inaugural World Inequality Report (2018) asserted the following:

"Economic inequality is widespread and to some extent inevitable. It is our belief, however, that if rising inequality is not properly monitored and addressed, it can lead to various sorts of political, economic, and social catastrophes."

Despite this, the theoretical consequences of inequality having societal effects are largely left unexplored. Economic models rarely consider inequality as an issue in and of itself. Most works are based on the assumption that inequality only affects individuals through either (i) their own consumption level, or (ii) indirectly through the social planner's equity or fairness principles.

That inequality itself might have an impact on the individual, either through simple preferences or societal effects, is not taken into account in this framework. Inequality is thus assumed impactless. Furthermore, this assumption is rarely explicitly stated or considered in-depth.

This work attempts to remedy this by treating inequality as an externality and considering the theoretical implications of such. This follows and extends on Thurow (1971), Alesina and Giuliano (2011), Rueda and Stegmueller (2016) and others. The work is divided into three main parts: *Framework* (Section 2), *Existence* (Section 3), and *Consequences* (Section 4).

In Section 2, a simple framework to consider the inequality externality in a welfarist context is introduced. This includes two other regularly used consequences of inequality, (i) the cumulative effect of diminishing marginal utility and (ii) generalized social weights. I describe their differences to the inequality externality intuitively and mathematically. Examples are used to illustrate that the inequality externality is distinct and potentially influential.

In Section 3, the existence of the inequality externality is examined. Can we, based on existing empirical evidence, reject a non-existence hypothesis? What are the difficulties in examining causal channels, and which are most likely to have an impact on the individual? The section contains several brief literature reviews and a discussion on the difficulties of causal empirical measurement. It concludes with an assessment of the likelihood of the externality's existence.

In Section 4, the consequences of the externality existing is examined. In particular, three pieces of economic theory are discussed. First, the failure of the First Welfare Theorem, from Thurow (1971). Second, the theoretical and numerical impacts on optimal nonlinear income taxation, from Støstad and Cowell (2019). Third, novel in this work, a discussion on commodity taxation and in particular the Atkinson-Stiglitz theorem, which no longer holds under an inequality externality.

Section 5 concludes.

2 Framework

The core idea of this thesis is that of an inequality externality. This corresponds to including inequality θ in the utility function $u(c_i, \theta, ...)$. Although the interpretation of this term is slightly different from the existing literature – see Section 2.1 – this inclusion is not in itself a novel concept. Thurow (1971), Carlsson et al. (2005) and Alesina and Giuliano (2011) have explored the same idea.

One novel part of the thesis, however, is the combination of the externality and other consequences of inequality. In particular, I argue that there are three main welfarist consequences of inequality, each distinct and corresponding to different intuitive causal channels:

1. The cumulative effect of diminishing marginal utility (DMU) of consumption
2. Generalized social weights
3. An inequality externality

Together these form a social welfare function;

$$SWF = \int \frac{g(c_i, X_i)}{u_{c_i}} u(c_i, \theta, X_i) di \tag{1}$$

Here c_i and θ in the utility function represents DMU and the inequality externality, respectively. $g(c_i, X_i)$ are the social weights.[1] The vector X_i represents individual parameters.

I will now briefly discuss the first two before dedicating more time to the externality.

Diminishing marginal utility: Much of the welfarist debate, and in particular the field of optimal taxation, uses diminishing marginal utility of consumption or income as a core assumption. This is introduced through a diminishing effect of c_i in $u(c_i, ...)$,[2] and is the essential driver towards more equality in a large span of economic models. In short: one dollar is worth less the more dollars you have.

Income is optimally allocated where the (social) marginal utility of consumption is highest, which is then at the bottom of the income distribution. The presence of income inequality runs contrary to this logic: Dalton (1920) summarizes the consequences as *"the extreme wastefulness from the point of view of economic welfare of large inequalities of income"*. This is the principal factor working in the direction of equality in large sections of economic literature.

[1] Saez and Stantcheva (2016) points out that their approach generalizes to the Pareto weights $\omega_i = g_i/u_{c_i}$, which is what I use here.

[2] Mathematically, $u_c > 0$ and $u_{cc} < 0$.

Generalized social weights: A second formalization of equality concerns are generalized social weights of the (simplified) form $g_i = g(c_i, X_i)$, as described in Saez and Stantcheva (2016). Here X_i are a set of personal characteristics determining whether society finds the individual more or less deserving of income: an example is the extent to which they were afforded equality of opportunity. If not, the argument goes, society should value them more. Given different weights, society cares more or less about inequality overall.

Saez and Stantcheva argue that these weights introduce value judgements into the welfarist debate, and can represent views corresponding to the Libertarian, Rawlsian, Utilitarian, concerns about equality of opportunity, and so on.

Importantly for this work, social weights are often used to include social planner-imposed redistributive justice. This is done by setting g_i equal to a value negatively proportional to u_i or the agent's income, which leads to more societal distaste for inequality. Saez (2001) writes that social weights in this context *"summarize in a transparent way the distributive objectives of the government"*.[3]

2.1 The inequality externality

Inequality entering the utility function in the form of θ in $u(c_i, \theta, X_i)$ complements this approach by considering efficiency concerns beyond consumption differences. This is the focus of Støstad and Cowell (2019) and other papers. In essence, this concerns the societal consequences of inequality. It includes any consequence of inequality that is individually felt by other agents than the low- or high-income agent.

The argument is simple; if inequality changes society in a way that affects utility, inequality should enter the utility function.

Inequality entering the individual utility can be appropriate in two different settings.

First, individuals may simply prefer more or less inequality. Living with high or low levels of inequality, or specific distributional features, may simply lower well-being. Altruism, jealousy or philosophical preferences may all contribute to this.[4] Note that these preferences may depend on the nature or causes of the inequality.

Second, inequality can lead to secondary effects such as social unrest or crime. If these *inequality effects* exist – if inequality changes one or more variables that changes utility – entering inequality itself into the utility function is an appropriate mathematical shorthand.

I also note one specific type of inequality effect: that of consumption. This implies that consumption levels are directly affected by inequality such that $c_i = c_i(\theta)$ (see Alesina and Giuliano (2011)).[5]

[3]Should preferences on inequality be introduced through the individual utility function or through social weights? The difference could correspond to the difference between individual happiness (which might be impacted by how society corresponds to your philosophy), and philosophical or political preferences. The former (preferences that impact your individual happiness) implies the utility channel. The latter (philosophical standpoints) implies social weights. Dividing the two is sensible if philosophical viewpoints can be fully independent from individual choice and happiness. This division is similar to one explored in Nyborg (2000).

[4]These preferences must impact the agent's utility, which could be an important distinction. We can imagine agents arguing for more or less equal distributions without it directly affecting their well-being. In this case social weights would be more appropriate (see footnote 3).

[5]Bergh et al. (2016) argues that this can be caused by purchasing power effects.

Briefly, then, inequality can affect (i) well-being directly through *preferences*, (ii) other variables that affect utility through *inequality effects*, here represented by $\Psi(\theta)$,[6] or (iii) consumption levels. We can mathematically formulate this, then simplify:

$$u(c_i(\theta), \theta, \Psi_i(\theta)) \rightarrow u(c_i, \theta) \tag{2}$$

The effect of θ is the net effect of the inequality externality.

The concept of an inequality externality has been notably explored by Thurow (1971) and Alesina and Giuliano (2011). The latter discusses utility functions and preferences for redistribution; the former is explored in Section 4.1. For more specific discussions, Rueda and Stegmueller (2016) (who look at inequality and crime) and Anbarci et al. (2009) (who look at inequality and traffic fatalities[7]) consider specific inequality effects in the theoretical context of inequality as an externality. Carlsson et al. (2005), and other papers by the same authors, explore many of the same concepts under the moniker inequality aversion.

It should be noted that these consequences are distinct from social weights. One can be a Libertarian who strongly believes that inequality has societal consequences, or one could be a Rawlsian who believes there are none.[8]

Overall these three factors – diminishing marginal utilities of consumption, generalized social weights, and the inequality externality – shape the principal's welfare optimization problem in Equation 1. I note that the first and third concepts are efficiency concerns. The second, from generalized social weights, is an equity concern.

Relativity concerns Relativity concerns, which are conceptually similar to inequality externality concerns, usually imply that an agents' income is a negative externality on other agents. In short: if the Joneses' income increases, your utility decreases.

Inequality concerns work in the same way for the top of the distribution, but not for the bottom. Essentially, now the Joneses' position in the income distribution matters. If the Joneses are top income earners, their increased income decreases other agents' utility, as with relativity concerns. But if the Joneses are at the *bottom* of the income distribution, their increased income *increases* other agents' utilities.

The difference can be put as this: Unlike with relativity concerns, an inequality externality constitutes a positive externality from bottom agents' income. Top incomes are a negative externality in both frameworks.

I note that relativity concerns can be more complex. Rank relativity concerns, for instance, cannot be societally decreased without perfect equality betweeen two or more agents, and have quite different policy implications.

Relativity concerns are firmly established as impactful in the empirical literature on subjective well-being. Results studying inequality concerns are more conflicting (see Senik (2005)). The connection between these two are seldom pointed out, however. This is an area of study that perhaps deserves more attention.

[6]Where Ψ can be a function of crime, health, and so on.

[7]Interestingly, there is a direct and physical connection between inequality and traffic fatalities. Vehicle size is correlated to income, and traffic fatalities are correlated to vehicle size disparity. Thus, inequality leads to traffic fatalities. Anbarci et al. (2009) examines this link empirically.

[8]See again footnote 3.

Many of the theoretical results I describe in Section 4 also hold for standard relativity concerns. The essential change – introducing externality concerns into what was previously only a self-selection problem – is the same. In fact, several of the papers I cite are based on relativity concerns, including Kanbur and Tuomala (2013) and Oswald (1983) (the groundwork for the analyses in Sections 4.2 and 4.3, respectively). In many cases the change is as simple as changing the weights in the externality.

2.2 Propositions

Using the framework from Equation 1 I can make two propositions.

PROPOSITION 1: *Different formalizations of the consequences of inequality are appropriate given different causal channels.*

Each formalization is inherently distinct, and appropriate with specific corresponding causal channels. Diminishing marginal utilities (formalization) will fail to capture the societal effects of inequality (causal channel), for instance.

For a sketch of a proof, see the following. DMU is taken into account in the individual agent's optimization problem, and skews the individual to prioritize leisure. Both other formulations leaves the individual's optimization problem unchanged, assuming the individual neglects their own impact on inequality.[9] However, the socially optimal work choice differs. With generalized social weights the individual makes the socially optimal work choice (absent taxes or other distortions). Under an inequality externality the individual makes a socially suboptimal work choice. Thus the three formalizations have distinct characteristics; which characteristics are appropriate depends on the model's causal channels.

As a result the optimal incentive structure offered by the social planner differs, and we have:

PROPOSITION 2: *Different formalizations of the consequences of inequality yield different policy implications.*

To illustrate I use a simple example from the optimal taxation literature.

A general result from the models based on Mirrlees (1971) is the (local) result of a zero marginal tax rate at the top (Sadka (1976) and Seade (1977)). This holds with any nonnegative generalized social weights, although introducing generalized social weights ordinarily changes the tax schedule in ways that cannot be captured by dimishing marginal utilities (Saez and Stantcheva (2016)). With an inequality externality, on the other hand, the top marginal tax rate is non-zero, as the income of the top earner presents a externality on other agents (Støstad and Cowell (2019)).The top rate is positive with a negative inequality externality, and negative with a positive inequality externality, which represents the internalization of the externality problem. Thus the three formalizations imply different optimal tax rates.

[9]Results hold without this assumption, which is made for simplification. If the individual takes their effect on inequality into account, their optimization problem changes marginally when introducing an inequality externality, as in a standard public good problem.

The proposition is not exclusive to optimal income taxation; it is a general result. It also implies that the exclusion of any of the three formalizations in general models is likely to yield misleading results.

Application: Reducing top incomes For a more illustrative example, imagine a society with N agents. One agent, Charles, has a vastly larger income than the $N-1$ other agents. Say we decrease Charles' income marginally. Now evaluate the welfare results from each formalization *alone* individually:[10]

1. DMU: $\Delta W < 0$. Welfare is unambiguously reduced when only considering the diminishing marginal utility of income. Charles' utility is lowered, and no other agents are affected.

2. Social weights: $\Delta W \leq 0$. Welfare is either reduced or unchanged. Charles' utility is decreased due to his lower income (a welfare loss), but this may be irrelevant if his social weight is zero. Unless Charles' social weight is negative there is no potential for welfare gain.[11]

3. Inequality externality: ΔW is ambiguous. Charles' utility is decreased from the income loss. On the other hand, assuming a negative inequality externality, every agent gains utility from the reduction in inequality. Total welfare change depends on the magnitude of the inequality externality.

So the three concepts have different implications. Additionally, the inequality externality introduces welfare changes that might seem reasonable in some contexts: the potential for a net welfare gain if an oligarch loses income, for instance.

The example uses that the marginal social welfare of income can be negative in the presence of distributional externalities, as pointed out by Carlsson et al. (2005) and Støstad and Cowell (2019).

3 Existence: Literature reviews

> *"What are the empirical effects of the income distribution on crime, social stability, political stability, or any other characteristic of society? Perhaps the impact is significant; perhaps it is insignificant. We just do not know."* - Thurow (1971)

Thurow wrote the above in 1971, and at the time it was undoubtedly true. There were essentially no empirical estimates of the effects of inequality on other characteristics, and even simple correlations were often unknown. Since then, however, the advent of big data sets and modern econometric techniques have made the study possible, if still difficult.

[10]Note that each formalization is evaluated in the presence of the standard case: a linear utility function and a utilitarian welfare function.

[11]Social weights are generally assumed to be positive. It is assumed that society never places negative weights on anyone's happiness.

In general, income inequality correlates with many markers of a good society; good health, low crime, high trust, and so on. These correlations, particularly on a cross-country level, are the essence of the popular book *The Spirit Level* (Wilkinson and Pickett (2009)). But the data may be misleading. Omitted variable bias, endogeneity and spurious correlations may all lead to correlation without causation.

In this section I take a critical look at the causational evidence for an inequality externality, or the idea that income inequality per se changes other societal factors. Overall, the section is meant to examine the *existence* of the inequality externality (through inequality effects particularly).

I will attempt to examine a simple hypothesis: income inequality, in itself, does not change other societal factors.

First I will discuss empirical difficulties specific to the problem, and what the ideal causal evidence would be. Second I conduct literature reviews on the existing evidence regarding the income inequality hypothesis on crime, health, political polarization, and trust, all of which could plausibly be argued to impact utility. Third I make an overall conclusion on the plausibility of the hypothesis.

3.1 Empirical difficulties in finding inequality effects

The study of inequality effects is rife with empirical difficulties. Below is a list of the most relevant.

Omitted variables: Any cross-sectional evidence will have issues with omitted variables. If a parameter differs between regions and affects both inequality and the outcome variable, it will lead to bias. A classical example is cultural differences between countries. As such, time-series data is generally required for causal statements.

Inaccurate or misleading data, and low coverage: Until recently, accurate and lengthy time-series data on inequality metrics were hard to come by. With recent efforts this is becoming less of an issue, although developing countries and pre-1950 data is still sparse.

Several outcome variables suffer from the same issue. This is often an intrinsic issue. Trust, for instance, is difficult to define, not to mention measure. Social unrest has the same issue. Deciding how to measure health outcomes and crime is controversial, as Bergh et al. (2016) and Rufrancos et al. (2013) respectively point out. Time-series data is often hard to come by.

Overall, these are common issues but often exacerbated with outcome variables such as trust, political polarization and social unrest.

Low variability in inequality over time: Most inequality metrics barely change over time. This, combined with the issues from cross-sectional data, makes identification doubly difficult. Effects must be identified through the effect of inequality changes over time, but these changes are very small.

Endogeneity: Most of the outcome variables – trust, health, crime and more – impact inequality themselves. This is a classic endogeneity problem, and could potentially be solved with exogenous variation. However, as we will see below, this is not always simple.

Exogenous variation: The above problems are common in empirical works. Identification strategies, as a response, use exogenous variation to identify the effect of only changing the control variable. But finding exogenous variation for inequality metrics is challenging at best, impossible at worst.

To illustrate: What changes inequality but is also uncorrelated to our outcome variables? It is a difficult question. Inequality has few determinants, and it is very difficult to argue that changes to these determinants (such as tax rates) are caused by exogenous change unrelated to outcome variables such as trust or health. The low variability of inequality metrics makes the issue even worse. This is a core difficulty in identifying inequality effects.

Non-linearities: When the World Inequality Report argues that inequality leads to "various sorts of political, economic, and social catastrophes", it is clear that they are to some extent discussing non-linear effects. This creates another challenge, particularly in first-difference models.

First difference models – often the best estimation of inequality effects we have – assume that the change in inequality with a time lag should correlate to the outcome variable (for instance social unrest). This requires a well-specified time lag *and* a well-defined functional form.

What if social unrest is low until a certain threshold, then increases drastically? Then first-difference models require very large amounts of data and may not be appropriate. Even worse, if the 'threshold' has a stochastic element in either time or magnitude, estimation is even further complicated.[12] This is not an unreasonable concern, and presents more empirical concerns.

Poverty and relativity effects Effects from absolute poverty and relativity can easily be picked up as inequality effects unless the empirical method is carefully specified. Pridemore (2011) argues that most of the existing evidence on crime and inequality is spurious due to the authors neglecting the effect of poverty; this issue also holds for other outcome variables. A similar argument is presented in Pare and Felson (2014), which finds a relationship in time-series data until poverty is controlled for.

The ecological fallacy: Group-level data Bergh et al. (2016) describes the ecological fallacy as what manifests when *"the correlations produced by aggregate observations are different from the underlying individual observations"*. They argue that group-level data is fundamentally unreliable when studying inequality effects.

[12]We can imagine a situation where inequality leads to social unrest, but only after inequality is above a certain threshold for a number of years. If inequality randomly decreases in the year before the social unrest breaks out, a simple first-difference model would show that more equality leads to social unrest.

Data on group averages alone cannot distinguish between an absolute income hypothesis, a relative income hypothesis, and an income inequality hypothesis. This is obviously problematic for the field, as the vast majority of papers use simple group averages. Individual-level data, where possible, is thus strongly preferred, and the value of group-level data is questionable.

Model Use: While not an empirical difficulty in itself, the econometric difficulty of the problem leads to many papers using the wrong econometric specification. This confounds the available evidence, particularly since different fields use different terminology, making it challenging to assert which papers are empirically sound.

To specify, many time-series papers include unjustified use of OLS models over either fixed/random effects (FEM, REM) or a general methods of moments (GMM) approach. However, even FEM and REM methods are potentially biased. To use inequality's effect on crime as an example, both FEM and REM methods are biased due to the persistence of crime over time. Meanwhile, the two-step GMM approach returns biased standard error estimates, as described in the original paper by Arellano and Bond (1991).[13] Overall, it is difficult to be econometrically accurate, and even moreso given the difficulties listed above.

3.1.1 The ideal empirical data

The econometric issues are clearly significant. Are they unavoidable? This section discusses what an ideal experiment to detect inequality effects may entail, in addition to individual-level data.

Exogenous fiscal variation Above I noted the dearth of exogenous variation. In special cases this may not apply. Mandated fiscal experiments from international organizations or specific global financial conditions *may* change a region's income distribution while being plausibly uncorrelated to other variables. Such an outside influence could function as an instrument, and with appropriate controls we could find the effect of inequality on other variables.

Randomized economic experiments on a smaller scale could have a similar but more precise effect. The trade-off would be more controlled circumstances for a loss of external validity. Such experiments would also be costly.

The smaller scale highlights a potential issue with the design. The *change* in income levels could lead to reactions based on how the distributional changes were put into place. Also, if inequality only has an effect through a significant time-lag or non-linearities are an issue, effects could be difficult to pick up. This represents a frustrating and recurring issue; even the best possible experiment would have significant flaws.

Other clever instruments might exist, but I do not know them.

Panel data The Arellano and Bond (1991) method – or, failing that, fixed effects or first differences[14] – is likely the best current method to find causality from inequality effects. The

[13] This bias can be fixed by a Windmeijer (2005) correction.

[14] These methods assume that the region and time fixed effects do not vary across time or individuals, respectively. Especially the first is troublesome in our setup. Also, if there are heterogeneous treatment

Arellano-Bond method essentially correlates lagged changes in inequality metrics to changes in the outcome variable, yielding consistent estimates with large samples.

There are still certain empirical issues, notably weak IV issues, already scarce variation being removed, non-linearities, and specifying the correct time lag. With a large and accurate data set, however, these could be minimized. Whether the current literature has sufficient data, and to what extent the assumptions of the method holds, is debatable. That being said, several papers claim to find significant results using the methods, and as of now they seem to be the best we have.

Lab experiments The most likely (and surprisingly least explored) source of data is lab experiments. While inequality *aversion* has been extensively studied in the lab, inequality *effects* have not, as far as I know.

Under controlled settings one could easily set up a treatment and control group with different levels of inequality and study the result on various outputs – cooperation, trust, output, "government" efficiency, and so on. There are, as far as I can see, two main issues with this approach.

The first issue is external validity, which is always an issue in laboratory experiments. It could be particularly true in this case, as it is easy to imagine experiment showing inequality effects that do not exist in real-world settings.

The second issue is the confounding of inequality effects and inequality aversion.[15] Both enter into the utility function as an inequality term, so both are part of the externality I discuss throughout the paper. However, if one wants to isolate inequality *effects*, as I do here, the agents' reactions to inequality can mask as an 'effect' and make this difficult. One solution would be to mask the income distribution from agents' knowledge. On the other hand, this could introduce ethical concerns.

If one could design such an experiment, an interesting test would be to see if the causal effect can be detected using standard methods (Arellano-Bond, first differences, and so on). It is quite conceivable that statistically different outcomes could be found in the control and treatment, thus proving an inequality effect, without these same effects being visible using standard methods (where variation is removed to deal with fixed effects and omitted variable issues). If the 'real' effect is large, such a non-detection with standard methods could be an interesting result indeed.

3.2 Literature reviews

This section is designed as a meta-analysis of the literature on various inequality effects. I will focus on existing literature reviews where possible; otherwise I will briefly discuss specific papers.

Ideally this section would also contain (i) theories of *how* inequality could cause change in the outcome variables,[16] and (ii) longer comments on specific individual papers. Due to

effects of inequality that are themselves correlated to the fixed effects, this presents another problem.

[15] In other words, preferences against inequality and actual society-wide effects one feels whether or not one dislikes inequality.

[16] As an example I will review how inequality can causally change crime rates in this footnote. There are three main theories cited in the literature.

time and length constraints this was not done.

In general I will not comment on correlations. However, it is worth noting that for most of the outcome variables, equality correlates with the better outcome; less crime, better health, more trust, and so on.[17]

3.2.1 Inequality and health

The literature on inequality's effect on health is perhaps the broadest of any effects; the topic has drawn interest within the fields of health economics, epidemiology, public policy and more.

The Bergh et al. (2016) book *Sick of Inequality?* is currently the most complete review of the academic literature. They largely comment on papers using individual-level data due to the ecological fallacy noted in Section 3.1, and find 87 such papers on various health topics.

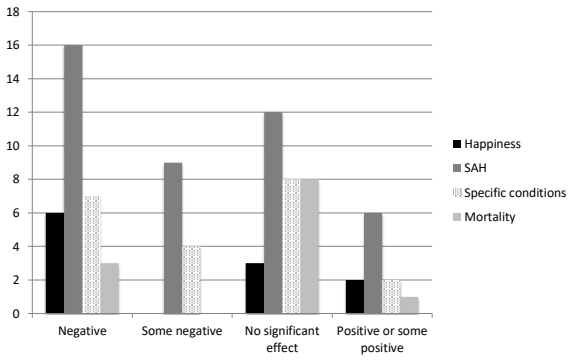


Figure 1: Figure 7.1 from Bergh et al. (2016) classifying the results of published academic works on inequality's effect on health.

Results are generally mixed, although there are some broader trends. Three main conclusions are drawn. First: 25 of 43 studies finds some evidence of a negative effect of inequality on self-assessed health. Second: There is little evidence of an inequality effect on mortality, with only three of twelve studies finding a negative effect (eight find no effect, one finds a positive effect). Third, studies on specific conditions (and happiness) are mixed, although a negative correlation is most common. Overall they argue that the link to self-assessed health is strong, and that the most likely causal channels are through political effects, stress, or trust.[18] They also

First there is the economic theory of crime, from Becker (1968). Becker theorizes that inequality leads to low-income agents, who have low returns to employment, close to high-income agents with valuable goods. The relative benefit of crime is thus increased, naturally increasing crime rates.

Second there is strain theory, from Merton (1938). This is, in some ways, similar to a relative income issue. If those around you exhibit high incomes, Merton argues, individuals feel frustrated at their position and are compelled to crime. Inequality increases this tension. Although this could be considered more a relative income-cost, inequality and relative income are closely linked (see Section 2.1).

Third, Shaw and McKay (1969) argues that crime follows from low social control and high disorganization, which often increases with general social heterogeneity and poverty, both connected to inequality. In other words, if inequality is high, social disorganization increases and thus crime increases.

The prohibitive length of a collection of these reviews should be clear. As a general comment, many works neglect to mention which causal channels they examine. Results and methods then become more difficult to analyze.

[17]See Wilkinson and Pickett (2009) and others for more. As an example of the general method to show associations, Kelly (2000) does a simple cross-sectional regression of violent crime rates on the Gini coefficient and other covariates in U.S. counties for 1991 data. The paper reports that the elasticity of the violent crime rate with regards to inequality is rather large at 1.3. There is no attempt at disentangling causation, but the paper concludes a "strong and robust impact" of inequality on violent crime.

[18]These are the authors' comments; few of the studied papers directly discuss the relevant causal channels.

Crime Type		Sig +ve assocs.	Sig -ve assocs.	No sig assocs.
All Crime		2		
Property Crime				
	Aggregate Property Crime	4	1	
	Auto Theft	2	1	
	Burglary	5		
	Shoplifting	1	1	
	Other Theft	1		
Violent Crime				
	Aggregate Violent Crime		3	
	Assault		2	
	Homicide and Murder	5	2	
	Rape		1	
	Robbery	3	1	1
	Serious Crime			1

Figure 2: The data from Rufrancos et al. (2013), indicating the time-series evidence from academic papers on inequality's effect on crime.

note, however, that the correlations might be driven by other variables such as poverty or relative income concerns.

Pickett and Wilkinson (2015) undertakes a study of the causal nature of the relationship, using an epidemiological framework of nine criteria to establish causality of a negative effect. Of these nine, the arguments for *Temporality*, *Cessation of exposure* and *Consideration of alternative explanations* arguments appear the weakest.[19] This is unsurprising, and in line with previous discussions. Regardless, the authors do present a large body of evidence and argue that the causal link is strong. They also point out that suicides have the opposite connection (indicating a *positive* inequality externality), and that increased social distances could be the most likely causal link.

Finally, Zheng (2012) reviews time-series papers on mortality with mixed results, and argues that not controlling for appropriate time lags distorts data. Using their own individual panel data they find no instantaneous effect, but significant effects with longer time lags (peaking at seven years).

3.2.2 Inequality and crime

Rufrancos et al. (2013) presents a systematic review of all pre-2013 time-series papers examining crime's relationship with inequality. They find 17 such papers, beginning in 1998, including several of the above. Ten are estimated with FEM, GMM or ARDL models (presumably the most reliable empirical methods).

Their conclusions are two-fold. First, there seems to be a strong specific effect of inequality on property crime such as burglary and car theft. This is in line with the economic theory of crime, wherein the relative gain of crime increases if inequality is high. However, the evidence on violent crime is weak and contradictory. The paper points out that this could be due to differences in the fraction of reported crimes, although it could also simply indicate a small or non-existent effect.[20] See Figure 2 for a summary of the results.

[19]The authors argue that they establish these links. However, many of the articles they cite as supporting evidence suffers from the flaws mentioned in Section 3.1. The paper is, in general, less rigorous with evidence than Bergh et al. (2016).

[20]Papers since their review offer some support for the association to violent crime; Menezes et al. (2013)

Top 1 Percent Income Share and Polarization in the US House of Representatives, 1913–2008

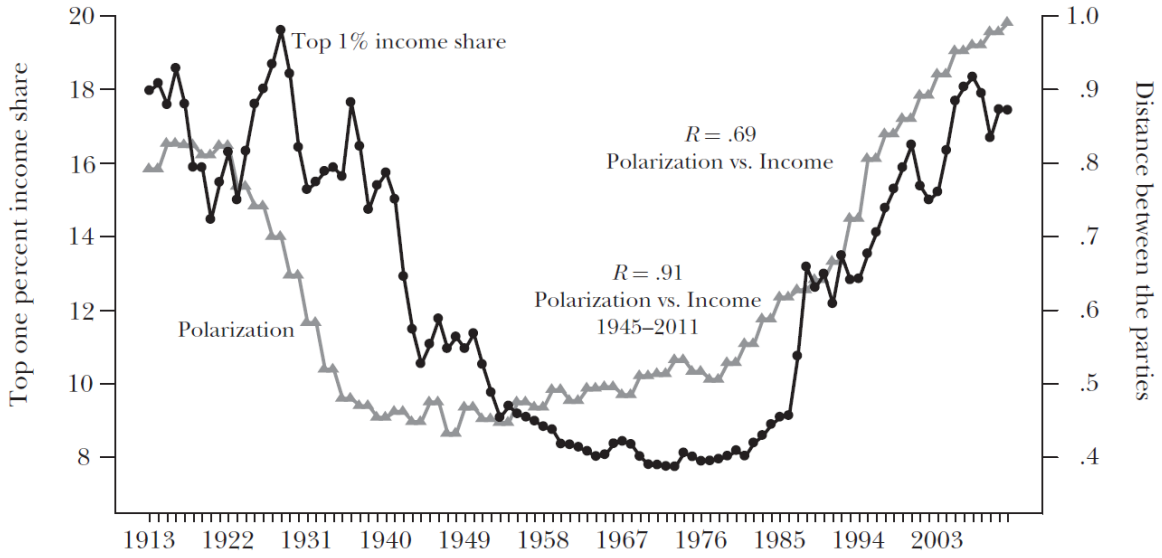


Figure 3: Figure from Bonica et al. (2013). Data from author's calculations on voting patterns in the U.S. House of Representatives and Piketty and Saez (2003).

3.2.3 Inequality and political polarization

The study of inequality and political polarization is, compared to the studies on crime and health, less developed. This seems to be largely due to a lack of convincing data on political polarization. The specific metric matters, and polarization is an inherently difficult concept to measure.

Of those who try, inequality is generally found to increase polarization. Bonica et al. (2013) convincingly argues that high income inequality has driven political polarization in the United States, emphasizing the key role of political donations (see Figure 3). Voorheis et al. (2015) use new state-level U.S. data to obtain what they argue is the first causal evidence on the link, noting a strong statistically significant connection between higher income inequality and more political polarization. Duca and Saving (2016) use longer time-series data to argue the same point. Finally, Pontusson and Rueda (2008) notes that the nature of the inequality may play a key determinant for the resulting polarization.

This study has grown in popularity given the recent rise in political polarization, particularly in the United States. Although empirical evidence is as of yet scarce, the few available papers are individually more convincing than many on health and crime.

and Enamorado et al. (2016) finds a positive association. On the other hand, Pare and Felson (2014) finds none and argues that the positive association is likely due to poverty being an omitted variable.

3.2.4 Inequality and trust

The correlation between high inequality and low trust in developed countries is strong – see, for instance, Elgar and Aitken (2010). The result does not necessarily indicate causation, however. Fairbrother and Martin (2013) find the same cross-sectional result in General Social Survey data, but are unable to detect a pattern over time. Gustavsson and Jordahl (2008) finds a negative effect in Swedish individual-level panel data. Finally, Barone and Mocetti (2016) finds a systematically negative relationship in developed countries in panel data, although some methods are dubious.[21]

Overall the empirical evidence on the topic is sparse and inconclusive, which presents ample room for more empirically rigorous research.

3.2.5 Other effects

This is a relatively short review of only four specific inequality effects. There are vigorous debates on inequality's effect on other variables, specifically growth and innovation, which I have avoided.[22] I note that these are more likely to yield a *positive* inequality externality. Other factors – such as civil unrest or the potential 'political, economic and social catastrophes' pointed out by the World Inequality Report are also not specifically considered due to the inherent data difficulties in exploring such effects.

3.3 Existence: Conclusion

The overall evidence on inequality effects is mixed. Empirical issues can affect even the most rigorous papers, and it is usually prohibitively difficult to conclusively establish causal channels. Almost any specific inequality effect may or may not exist. Causation is simply difficult to assert, despite a multitude of papers on each subject.

However, due to the sizable number of *plausible* hypotheses it is perhaps even more challenging to argue that there are *no* effects of inequality. Some effects, such as those of inequality on property crime and self-assessed health, are independently strong. Other plausible connections, in particular that of political polarization, are left inconclusive largely due to a lack of data. It is difficult to argue that all the available evidence is spurious.

In conclusion, the distributional nature of a society seems to have an impact on the agents in that society. The size and specific channels of these effects may be almost impossible to pin down. However, the overall existence of an inequality externality seems probable, or even likely.

This sets up my next discussion: on the consequences of such an externality.

[21] This is particularly true for their three instruments for inequality: (i) the initial sector composition (2-digit) of the country, (ii) technological coefficients capturing the sectoral dependence on ICT and (iii) the worldwide growth of the ICT industry. While they find the same result with fixed effects, these instruments are unlikely to be valid.

[22] Primarily because the channel from growth and innovation to utility is strongly contingent on an intertemporal dimension, confusing the framework.

4 Consequences: Theory

In the previous section I discussed the existence of the inequality externality. In this section I take the existence for granted, and look for theoretical results given this assumption.

The essential driver of the theoretical changes is the new nature of the problem. Stiglitz (1982) described a large class of models, including the optimal taxation problem, as pure self-selection problems. This changes when inequality enters the utility function; then the problem is both a self-selection *and* an externality problem, which leads to substantially different conclusions.

4.1 The First Welfare Theorem (Thurow, 1971)

Thurow (1971) introduced the concept of inequality as a public good. The paper is a short one, wherein Thurow points out two reasons for why the income distribution might be an argument in the individual's utility function: (i) agents might have a taste for redistribution, and (ii) there might be externalities associated with the distribution of income. He considers a similar notation as used in this paper, where individual utility $u(Y_i, X)$ is composed of individual income Y_i and the income distribution X.

From this approach, Thurow notes that not every initial distribution is Pareto optimal. His conclusion is that redistribution might be necessary to reach a Pareto optimum. While he does not phrase it as such, the implication is the failure of the First Welfare Theorem.[23]

This implies that redistribution can make *everyone* better off. This argument, often presented by those in favor of redistribution, had little to no basis in economic theory before Thurow. This illustrates how the inequality externality can help unify theory and public intuition.

4.2 Optimal income taxation (Støstad and Cowell, 2019)

Støstad and Cowell (2019), currently a working paper, discusses the introduction of an inequality externality into the Mirrlees (1971) optimal nonlinear income taxation model.

When the Mirrlees model was first developed, few theoretical conclusions could be found. The Mirrlees model is highly non-specific, wherein the incentive to raise taxes to increase government revenue to redistribute is weighed up against the incentive to lower taxes to increase work effort and thus output. Due to the generality of the problem, few clear results were initially found without imposing restrictive structure on the problem.

Eventually, however, some specific points were made. Sadka (1976) and Seade (1977) in particular found important results, namely:

1. The optimal marginal tax rate at the top is zero, assuming a bounded skill distribution.
2. The optimal marginal tax rate at the bottom is zero, assuming everyone works.
3. The optimal marginal tax rate is between zero and one.

[23] Given that the income distribution is considered a public good and thus an externality, this is not a surprising result: the FWT explicitly assumes no externalities.

Other works also made the theory more tractable, and in general more academically interesting. Diamond (1998) showed that a set of simple conditions were enough to ensure a U-shaped marginal tax structure. Saez (2001) noted that simple labour supply elasticities were sufficient to deduce optimal marginal tax rates, assuming quasi-linear utility. It was empirically found that most income distributions follow a Pareto distribution near the top, allowing simplification of the optimal marginal tax rate formulas.

In Støstad and Cowell (2019), the standard theory is modified by introducing an additive inequality function $\Gamma(\theta)$ into individual utility:

$$U(c, h, \theta) = u(c) - V(h) - \Gamma(\theta) \tag{3}$$

Where inequality is measured with the Gini, $\theta = \int \kappa(n)x(n)dG(x)$, where $\kappa(n) = \frac{2F(n)-1}{\mu}$ is the weight of the agent in the Gini (negative at the bottom and positive at the top).[24] The optimal marginal tax rates are then:

$$\frac{t}{1-t} = \frac{\gamma}{\lambda}\left[\kappa(n) + \frac{\zeta}{f(n)n}\int_n^\infty \frac{u_{x(n)}}{u_{x(p)}}\kappa(p)f(p)dp\right] + t_{orig} \tag{4}$$

Where t_{orig} are the tax rates from the no-externality model.

$\frac{\gamma}{\lambda}$ is the shadow price of inequality in units of public goods.[25] $\zeta = \frac{V_{hh}h}{V_h} + 1$ is closely related to the inverse compensated elasticity of labor,[26] and n is intrinsic earning ability. In discussing the results a negative inequality externality will be assumed, i.e. a positive $\frac{\gamma}{\lambda}$.

There are two new terms. The first internalizes the individual's externality, decreasing and increasing marginal tax rates at the bottom and top respectively. The second, inside the integral, represents the social planner's increased willingness to decrease inequality. An increase in marginal rates decreases the income of every agent above the tax bracket, which decreases inequality; the social planner thus has an additional incentive for higher tax rates.

Using this result, all the theoretical points above fall away. Instead the following results emerge:

1. Assuming a negative/positive inequality externality, optimal marginal tax rates at the top are positive/negative.

2. Assuming a negative/positive inequality externality and that everyone works, optimal marginal tax rates at the bottom are negative/positive.

3. Optimal marginal tax rates are bounded below one, but can be negative.

Overall, tax rates are unambiguously more progressive (regressive) assuming a negative (positive) inequality externality.

An additional point often made in the debate around optimal taxation is that of the Laffer rate, or revenue-maximizing taxation. In general it has been considered a simple

[24] This formulation, and the rest of the paper, currently assumes second-order conditions hold. Proving this is a work in progress.

[25] The full expression is $\frac{\gamma}{\lambda} = \frac{\int_0^\infty \frac{\Gamma_\theta}{u_x} f(n)dn}{1-\int_0^\infty \frac{\Gamma_\theta}{u_x}\kappa(n)f(n)dn}$.

[26] With quasi-linear utility, $\zeta = \frac{1}{E_c} + 1$. Also in this case, $u_x = 1$.

upper bound on optimal tax rates. As an example, Piketty et al. (2014) writes that they "focused on the revenue maximizing top tax rate, which provides an upper bound on top tax rates".

When introducing an inequality externality, top marginal tax rates above the Laffer rate are possible and often expected. If one discounts the welfare of those at the top[27] it is in fact a mechanical result for top-earners.[28] In any such case, *any* additive negative inequality externality leads to a top marginal tax rate above the revenue-maximizing rate. If inequality is an externality, the Laffer rate is not a logical bound; it is simply the rate at which we maximize revenue.

Many of these findings can be summarized as follows. Usually optimal taxation explores the optimal balance of three effects: the mechanical, behavioral and welfare effects of taxation. By introducing inequality concerns, one must also take into account the *equality* effects of taxation. In short, equality is a benefit or ill in itself, and thus changing it also changes welfare.

This explains why the top marginal tax might be positive, for instance. Although there is no revenue to be gained by raising the top tax rate, there is a positive equality effect. If only the welfare effect is taken into account, the rate should be zero – introducing the equality effect increases the top tax rate to a positive number.

Numerical estimates support the theoretical claims. In addition, they indicate that the changes in the optimal rates might be considerable.

The following individual utility function was used for the numerical simulations:

$$U(x,h,\theta) = x - \frac{h^{\left(1+\frac{1}{E_c}\right)}}{\left(1+\frac{1}{E_c}\right)} - \eta\mu\theta \tag{5}$$

This specification keeps disutility from inequality proportionally constant to utility from consumption over any chosen unit of x.[29] The optimal tax rates thus specified are,

$$\frac{t}{1-t} = \eta(2F(n)-1) + \frac{1-F(n)}{f(n)n}\left(\frac{1}{E_c}+1\right)(1+\eta F(n)) \tag{6}$$

Where E_c is the compensated elasticity of labor supply. These simulations are done for the simple cases of (a) a full Pareto productivity distribution with $a = 2$,[30] and (b) a full lognormal distribution with $\sigma = 0.39$ and $\mu_{log} = -1$.[31] In both cases $E_c = 0.3$. The optimal marginal tax rates for a selection of η are displayed in Figure 4.

The various levels of η are set by the magnitude of the inequality externality. In lieu of accurate macroeconomic estimates, a simple hypothetical give them meaning: What multiple of their current income should the average agent require to move from Denmark-like to United

[27]This is done in a Rawlsian utility function, for instance, where the welfare of the lowest-welfare agent is maximized, or in any SWF that approaches a zero weight as utility/income increases.

[28]And similarly at the bottom in the case of a positive inequality externality given that everyone works.

[29]Due to the presence of μ in the inequality term, which is equivalent to using the absolute Gini index, as is done in Section 4.3.

[30]Using Saez (2001)'s estimation for the top of the U.S. income distribution.

[31]The standard simulation numbers for this literature, see Tuomala et al. (1990).

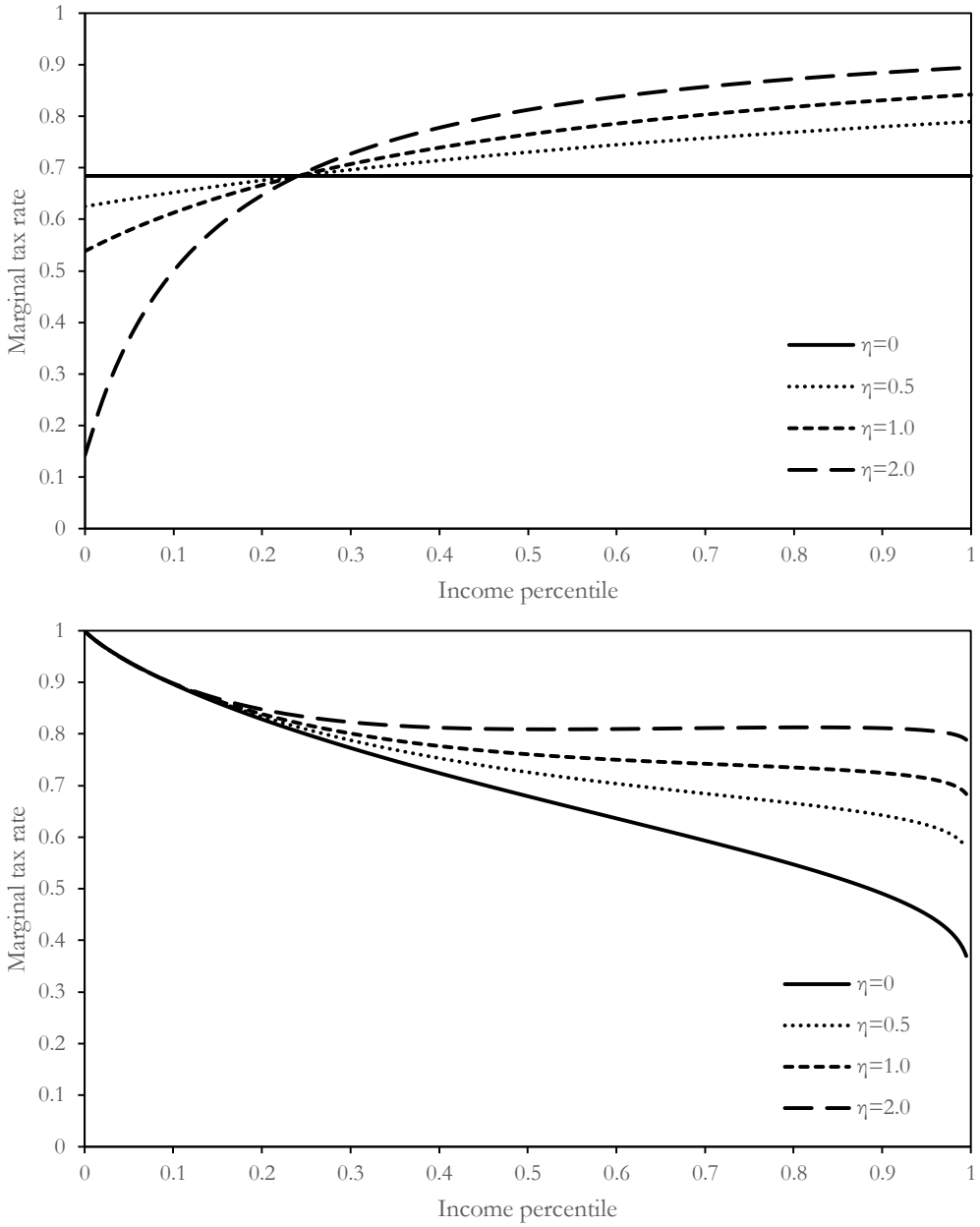


Figure 4: Optimal marginal tax rates for various negative inequality externality magnitudes η for two different ability distributions. The social planner is Rawlsian and the productivity distribution is (a) a Pareto distribution with a=2, (b) a log-normal distribution with $\sigma = 0.39$ and $\mu_{log} = -1$. See Table 1 for an explanation of the inequality externality magnitudes. Notably, $\eta = 0$, the solid line, is the standard case of no inequality externality. $\eta = 2.0$ implies that a representative agent with mean income in a society with Denmark-like income inequality would be indifferent to increasing her income by 25% at the same time as income inequality increased to the United States' current income inequality level.

	$\eta = 0.0$	$\eta = 0.5$	$\eta = 1.0$	$\eta = 2.0$
Equivalent income ratio	1	1.06	1.13	1.25

Table 1: The equivalent income ratios in societies with Denmark- and United States-like inequality for various levels of the inequality externality η (from Equation 5). As an example: the equivalent income in the U.S. society for an average agent in the Denmark society when $\eta = 0.5$ is 1.06 times the average Danish income.

States-like inequality? These equivalent incomes and their corresponding η are shown in Table 1.[32]

Here I also use data from Carlsson et al. (2005), a survey experiment to find macroeconomic inequality aversion estimates, to approximate values of η. This data and result was intended to be a bigger part of the thesis, comparing utility function fits, but due to methodological constraints it was not possible to find results for η without arbitrarily constraining the functional form. The median value in the experiment, when the utility function is constrained to Equation 5, is approximately $\eta = 1.00$. A majority of respondents have $0.26 < \eta < 2.18$.[33]

The graphical representation makes the increased progressivity of the resulting tax rates obvious, particularly for large η. The introduction of even a small inequality externality substantially changes the optimal tax structure.

The high marginal tax rates themselves should not be surprising, given that a Rawlsian social planner and a negative inequality externality was assumed. Instead, the large magnitude of the change from the standard Mirrlees case ($\eta = 0$) is notable.

The top marginal tax rate in the Pareto-distributed case increases from 68% to 89% when assuming an equivalent income of 1.25μ ($\eta = 2.0$).[34] In the log-normal case the top marginal tax rate increases from 0% to 67%.[35] These changes are large, and illustrate the potentially significant impact of the inequality externality.

In conclusion both theoretical and numerical results from the optimal income taxation model change substantially under an inequality externality, which emphasizes its potential importance.

[32] Assuming the same leisure, that the mean income difference between the two countries is negligible, and that relative position is irrelevant; according to the 2017 World Economic Outlook database GDP per capita is \$61803 in Denmark, and \$59707 in the United States. Calculations are based on Gini coefficients of 0.410 for the United States and 0.285 for Denmark.

[33] The frequency distribution is the following (cumulative frequency in parenthesis): $-\infty < \eta \leq -0.89 = 0.03\,(0.03)$, $-0.89 < \eta \leq 0.00 = 0.04\,(0.07)$, $0.00 < \eta \leq 0.26 = 0.11\,(0.18)$, $0.26 < \eta \leq 0.51 = 0.11\,(0.29)$, $0.51 < \eta \leq 1.11 = 0.23\,(0.52)$, $1.11 < \eta \leq 1.56 = 0.21\,(0.73)$, $1.56 < \eta \leq 2.18 = 0.11\,(0.84)$, $2.18 < \eta \leq 2.53 = 0.07\,(0.90)$, $2.53 < \eta \leq 3.00 = 0.04\,(0.94)$, $3.00 < \eta = 0.06\,(1.00)$. Given that we constrain the utility function to our functional form, these numbers should be interpreted with extreme caution.

[34] The Pareto-distributed simulations are more accurate towards the tail of the income distribution (where the actual income distribution begins to approximate a Pareto tail). Also note that the zero-at-the-top result does not appear in the standard case as the Pareto distribution is unbounded.

[35] This illustrates the Pigouvian correction at the top. The very local zero tax at the top-result is not visible in the graph, but is borne through in the simulations. At the 99th percentile the marginal tax rate increases from 39% to 79%.

4.3 The Atkinson-Stiglitz theorem

The Atkinson-Stiglitz theorem, from Atkinson and Stiglitz (1976), states the following:

If the utility function is weakly separable between labor and all consumption goods (taken together), then no commodity taxation need to be employed ($\tau_i = 0$).

In this section I intend to show that the theorem fails when an inequality externality is introduced to the individual's utility function.

The problem is a modification of the original Atkinson and Stiglitz (1976) problem, and similar to Mirrlees (1976) and equivalent to a special case of Oswald (1983). The technical contribution of this work is thus small.

We want to design a tax system to maximize welfare from a vector of goods x, a numeraire good z (leisure), inequality θ, prices p and earnings ability n:

$$W = \int u(x, z, \theta, n) f(n) dn \tag{7}$$

The main addition is considering that utility is directly affected by θ, an inequality measure. We use the (absolute) Gini:

$$\theta = \int \kappa(n) p \cdot x(n) f(n) dn \tag{8}$$

$\kappa(n)$ is the weight of the agent in the Gini, $\kappa(n) = 2F(n) - 1$.[36] The absolute Gini is the standard Gini multiplied by mean income, and it is used to achieve scale invariance in the trade-off with consumption. $p \cdot x(n)$ is a scalar equivalent to the agent's total income (consumption). Note that we ignore the numeraire good z here, i.e. leisure is not included in the inequality term.

The formulation does not necessarily imply that agents actively dislike inequality per se. It may equally well imply that agents' utility is exogenously affected by inequality through secondary variables such as crime, political distortions, or health effects, which are in turn causally changed by inequality.

We approach in a similar fashion to Oswald (1983). There are, in total, three constraints.

First, individuals maximize their own consumption. Individuals do not take their effect on inequality into account – an assumption equivalent to there being a continuum of agents. This, together with second-order conditions holding by assumption, leads to the constraint,

$$u_x x'(n) + u_z z'(n) = 0$$

The second constraint is the production constraint:

$$A \geq \int [px(n) + z(n)] f(n) dn$$

Third, the inequality constraint from Equation 8.

[36] Note for further study: Using the agent's weight in the skill distribution assumes that second-order conditions hold.

That leads to the full maximization problem; choosing the optimal non-linear tax schedule such that the following problem is solved:

$$max \quad W = \int u(x(n), z(n), \theta, n) f(n) dn$$

$$A - \int [p \cdot x(n) + z(n)] f(n) dn \geq 0$$

$$\theta - \int \kappa(n) p \cdot x(n) f(n) dn = 0$$

$$u_x x'(n) + u_z z'(n) = 0$$

I create a Lagrangean:

$$L = \int \{[u(x(n), z(n), \theta, n) - \lambda(p \cdot x(n) + z(n)) - \gamma \kappa(n) p \cdot x] f \tag{9}$$

$$+\lambda A + \gamma \theta + \mu(u_x x'(n) + u_z z'(n)) dn \tag{10}$$

We integrate by parts and find the first-order conditions (neglecting corners):[37]

$$x : [u_x - \lambda p - \gamma \kappa p] f - \mu'(n) u_x - \mu u_{nx} = 0 \tag{11}$$

$$z : [u_z - \lambda] f - \mu'(n) u_z - \mu u_{nz} = 0 \tag{12}$$

$$\theta : \int \Big(u_\theta f + \mu[u_{x\theta} x'(n) + u_{z\theta} z'(n)] \Big) dn + \gamma = 0 \tag{13}$$

$$\mu(0) = \mu(\infty) = 0 \tag{14}$$

The main result is found through combining Equations 11 and 12:[38]

$$t(n) = s(n) - p = \frac{\mu(n) u_z s_n}{\lambda f(n)} + \frac{\gamma \kappa(n) p}{\lambda} \tag{15}$$

Again identical to the Oswald (1983) result. s is defined as $s = \frac{u_x}{u_z}$, and $u_{nx} - s u_{nz} = u_z s_n$ was used in the derivation. This equation is the standard Atkinson and Stiglitz (1976) and Mirrlees (1976) result with an additional second term.

This gives the impression of strict additivity, which can be misleading if γ has a complicated form (see Equation 13). If $u_{x\theta} = u_{z\theta} = 0$, however, or we do have separability in the utility function, then the new tax function is indeed simple.

[37] The first-order conditions for x and z use $\mu u_n + \mu' u$ in place of $\mu(u_x x'(n) + u_z z'(n))$. To see this we note that the IC is equivalent to $v'(n, I) = u_n(x(n), z(n), I, n)$ (defining $v(n, I) = u(x(n), z(n), \theta, n)$). Through integration by parts and substitution, $\mu v'(n, I) = \mu' u(x(n), z(n), \theta, n)$.

[38] Solving for $\mu'(n)$ and substituting in quickly yields the result.

Now to the case of the Atkinson-Stiglitz theorem. If utility is separable between leisure z and all commodities x, then s_n is zero. Without our new second term, that implies $\tau = s - p = 0$ for all goods, or no commodity taxation. This is the standard version of the theorem.

However, in our case this is not true. Even if $s_n = 0$, the inequality modifier $\frac{\gamma\kappa(n)p}{\lambda}$ implies differential tax rates.

Let us investigate the sign of this term $\frac{\gamma\kappa(n)p}{\lambda}$. A negative inequality externality generally leads to a positive γ.[39] λ is also positive from KKT conditions. p is by definition positive. And finally, $\kappa(n)$ is the weight of the agent in the Gini, negative at the bottom and positive at the top.

Overall, then, the term is negative for n below the mean, and positive for n above the mean.

Optimal marginal tax rates are thus generally ambiguous, as we do not know ability, and have to set the same tax rate for every consumer. However, presumably goods primarily used by the poor should be subsidized, and goods primarily used by the rich should be taxed. More work on this is required.

The result implies that a second instrument is needed to optimize tax schedules. This is a somewhat puzzling result. Redistribution is still the primary goal inherent in the problem, although there are now three redistributive goals – to gather revenue, to reduce inequality, and to internalize the externality problem. One would think all forms of redistribution could be achieved through the income tax, but this is mathematically untrue. The exact reason for this is unknown, but I hypothesize that the second instrument may be necessary to get agents to internalize the inequality externality. Again, more work is necessary.

The Atkinson-Stiglitz theorem is broadly used in policy debate, and is often invoked in proposals for commodity taxation reform. It is noteworthy that it falls away when inequality is considered an externality, and again shows that the assumption that inequality has no societal effects is a significant one.

The result is particularly remarkable given the somewhat unintuitive nature of the original theorem (that some channels of redistribution are left unused). If the unintuitive result is born not out of intrinsic truth but instead out of a lack of model complexity, the theorem itself may mislead more than inform.

As a broader point, many famous theoretical works in Economics are based on elegant yet fragile mathematics. That these results are vulnerable to externalities is well-known. When considering inequality as an externality, that vulnerability takes on added significance.

Finally, I note that this and the optimal income taxation results strengthen Propositions 1 and 2 from Section 2; different types of inequality concerns imply different policy interventions.

[39] For the case of an additive inequality externality, this can be seen in Equation 13 (as then $u_{x\theta} = u_{z\theta} = 0$). It is also an intuitive result: if inequality is a negative externality, the shadow price of the inequality constraint should be positive.

5 Conclusion

Economic inequality has consequences for the individual beyond their own level of consumption. These consequences, when introduced into economic models, substantially changes traditional results. Using an inequality externality helps realign economic theory with public policy and rhetoric.

I have argued the above through three individual arguments.

First: The inequality externality is intuitively and mathematically distinct from other methods of introducing inequality concerns.

To illustrate why this is important, note that the externality alone models the consequences of inequality's societal effects. If the externality is mathematically distinct from other standard methods – namely diminishing marginal utilities of income and social weights – it follows that *standard models assume that inequality has no societal effects.*

Second: Based on the current empirical evidence, inequality is likely have societal effects.

It is theorized that inequality changes many variables that matter for our well-being. However, conclusively proving these theories is challenging due to intrinsic empirical difficulties. By reviewing the literature on inequality's effects on health, crime, trust, and political polarization, it was found likely that inequality is very likely to have at least *some* effect on variables that change well-being. As such, the inequality externality likely exists. Its magnitude is difficult to ascertain, an issue unlikely to be resolved in the near future.

Third: Introducing an inequality externality into economic models substantially changes theoretical and numerical results.

To show this three different theoretical results were emphasized: (i) the failure of the First Welfare Theorem, (ii) the large theoretical and numerical effects on nonlinear optimal income taxation, and (iii) the failure of the Atkinson-Stiglitz theorem. These are results with wide-ranging implications for tax reform and the future of public redistributive policy.

In short; many of the models we consider canonical have assumed away inequality concerns. The resulting conclusions must be considered in that light. When that assumption is removed, many theoretical results fall away, leaving a host of new conclusions.

Weaknesses of this model include the reliance on specific functional forms – empirical difficulties prevents us from knowing the actual impact of inequality on other factors – and the lack of more immediately policy-relevant conclusions. To some extent these difficulties may be ameliorated by more research.

Finally, the above indicates that the exclusion of an inequality externality is an issue in specific theoretical models. It is worth noting that the same argument holds for broader empirical models intended for policy analysis; not including an inequality externality often distorts results and sets the costs of inequality at artificially low levels.

Further work

The two primary directions of further work is to (i) understand the magnitude and functional form of the inequality externality, and (ii) further examine theoretical consequences, particularly in models used for policy analysis.

For a specific project, a laboratory experiment showing inequality effects that cannot

be detected with traditional methods (see Section 3.1.1) could help our understanding for why inequality effects are commonly referred to in public rhetoric, yet rarely discussed in economic theory.

Acknowledgements

I would like to thank Daniel Waldenström, who supervised the project and provided invaluable comments and advice. I am also grateful to Stéphane Gauthier, Frank Cowell, and Andrew Clark, who helped me with the idea and specific theoretical models, and Fredrik Carlsson and Olof Johansson-Stenman, who shared their data and discussed various topics. Finally, Marie Young Brun, Paula Sanchez, Simon Altman, Kieran Byrne, and Elif Cansu Akoğuz gave me their time and thoughts, for which I am grateful. All remaining errors are mine.

References

Alesina, A. and Giuliano, P. (2011). Preferences for redistribution. In *Handbook of social economics*, volume 1, pages 93–131. Elsevier.

Anbarci, N., Escaleras, M., and Register, C. A. (2009). Traffic fatalities: does income inequality create an externality? *Canadian Journal of Economics/Revue canadienne d'économique*, 42(1):244–266.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of public Economics*, 6(1-2):55–75.

Barone, G. and Mocetti, S. (2016). Inequality and trust: new evidence from panel data. *Economic Inquiry*, 54(2):794–809.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime*, pages 13–68. Springer.

Bergh, A., Nilsson, T., and Waldenström, D. (2016). *Sick of Inequality?: An Introduction to the Relationship Between Inequality and Health*. Edward Elgar Publishing.

Bonica, A., McCarty, N., Poole, K. T., and Rosenthal, H. (2013). Why hasn't democracy slowed rising inequality? *Journal of Economic Perspectives*, 27(3):103–24.

Carlsson, F., Daruvala, D., and Johansson-Stenman, O. (2005). Are people inequality-averse, or just risk-averse? *Economica*, 72(287):375–396.

Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal*, 30(119):348–361.

Diamond, P. A. (1998). Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95.

Duca, J. V. and Saving, J. L. (2016). Income inequality and political polarization: time series evidence over nine decades. *Review of Income and Wealth*, 62(3):445–466.

Elgar, F. J. and Aitken, N. (2010). Income inequality, trust and homicide in 33 countries. *European Journal of Public Health*, 21(2):241–246.

Enamorado, T., López-Calva, L. F., Rodríguez-Castelán, C., and Winkler, H. (2016). Income inequality and violent crime: Evidence from mexico's drug war. *Journal of Development Economics*, 120:128–143.

Fairbrother, M. and Martin, I. W. (2013). Does inequality erode social trust? results from multilevel models of us states and counties. *Social science research*, 42(2):347–360.

Gustavsson, M. and Jordahl, H. (2008). Inequality and trust in sweden: Some inequalities are more harmful than others. *Journal of Public Economics*, 92(1-2):348–365.

Kanbur, R. and Tuomala, M. (2013). Relativity, inequality, and optimal nonlinear income taxation. *International Economic Review*, 54(4):1199–1217.

Kelly, M. (2000). Inequality and crime. *Review of economics and Statistics*, 82(4):530–539.

Menezes, T., Silveira-Neto, R., Monteiro, C., and Ratton, J. L. (2013). Spatial correlation between homicide rates and inequality: Evidence from urban neighborhoods. *Economics Letters*, 120(1):97–99.

Merton, R. K. (1938). Social structure and anomie. *American sociological review*, 3(5):672–682.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208.

Mirrlees, J. A. (1976). Optimal tax theory: A synthesis. *Journal of public Economics*, 6(4):327–358.

Nyborg, K. (2000). Homo economicus and homo politicus: interpretation and aggregation of environmental values. *Journal of Economic Behavior & Organization*, 42(3):305–322.

Oswald, A. J. (1983). Altruism, jealousy and the theory of optimal non-linear taxation. *Journal of Public Economics*, 20(1):77–87.

Pare, P.-P. and Felson, R. (2014). Income inequality, poverty and crime across nations. *The British Journal of Sociology*, 65(3):434–458.

Pickett, K. E. and Wilkinson, R. G. (2015). Income inequality and health: a causal review. *Social science and medicine*, 128:316–326.

Piketty, T. and Saez, E. (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics*, 118(1):1–41.

Piketty, T., Saez, E., and Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American economic journal: economic policy*, 6(1):230–71.

Pontusson, J. and Rueda, D. (2008). Inequality as a source of political polarization: A comparative analysis of twelve oecd countries. *Democracy, inequality, and representation*, pages 312–353.

Pridemore, W. A. (2011). Poverty matters: A reassessment of the inequality–homicide relationship in cross-national studies. *The British Journal of Criminology*, 51(5):739–772.

Rueda, D. and Stegmueller, D. (2016). The externalities of inequality: Fear of crime and preferences for redistribution in western europe. *American Journal of Political Science*, 60(2):472–489.

Rufrancos, H., Power, M., Pickett, K. E., and Wilkinson, R. (2013). Income inequality and crime: A review and explanation of the timeâ series evidence. *Sociology and Criminology-Open Access*.

Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *The review of economic studies*, 43(2):261–267.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229.

Saez, E. and Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(1):24–45.

Seade, J. K. (1977). On the shape of optimal tax schedules. *Journal of public Economics*, 7(2):203–235.

Senik, C. (2005). Income distribution and well-being: what can we learn from subjective data? *Journal of Economic Surveys*, 19(1):43–63.

Shaw, C. R. and McKay, H. D. (1969). Juvenile delinquency and urban areas: A study of rates of delinquency in relation to differential characteristics of local communities in american cities (1969).

Stiglitz, J. E. (1982). Self-selection and pareto efficient taxation. *Journal of public economics*, 17(2):213–240.

Støstad, M. N. and Cowell, F. (2019). The inequality externality and the optimal nonlinear income taxation model. *Working paper*.

Thurow, L. C. (1971). The income distribution as a pure public good. *The Quarterly Journal of Economics*, pages 327–336.

Tuomala, M. et al. (1990). Optimal income tax and redistribution. *OUP Catalogue*.

Voorheis, J., McCarty, N., and Shor, B. (2015). Unequal incomes, ideology and gridlock: How rising inequality increases political polarization. *Ideology and Gridlock: How Rising Inequality Increases Political Polarization (August 21, 2015)*.

Wilkinson, R. and Pickett, K. (2009). *The Spirit Level: Why More Equal Societies Almost Always Do Better*. Allen Lane.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of econometrics*, 126(1):25–51.

Zheng, H. (2012). Do people die from income inequality of a decade ago? *Social Science and Medicine*, 75(1):36 – 45.