Exploitation: theory and empirics*

Daniele Girardi[†] Nicolas Grau[‡] Roberto Veneziani[§] Naoki Yoshihara[¶]

April 16, 2025

Abstract

This paper provides a novel axiomatic analysis of exploitation as the unequal exchange of labour, derives an empirical exploitation index at the individual level, and estimates its distribution in the US in 1975-2022. We show that, among possible definitions of exploitation, only one satisfies a small set of formally weak and normatively salient axioms. From this definition, we derive an individual-level exploitation intensity index which provides a new measure of well-being and inequality, complementary to existing ones and able to jointly take into account the distributions of income and work time. In US data, exploitation intensity provides additional information compared with standard income inequality measures and predicts important well-being and political outcomes. Inequality in exploitation increased more than income inequality since 1975.

JEL: D63 (Equity, Justice, Inequality, and Other Normative Criteria and Measurement); D5 (General Equilibrium and Disequilibrium).

*We are grateful to Duncan Foley and John Roemer for enlightening discussions that have significantly contributed to shape the main ideas of this paper. We thank Andrea Attar, Meghnad Desai, Peter Flaschel, Marco Mariotti, Peter H. Matthews, Dan McGee, Simon Mohun, Juan Moreno Ternero, Marco Ranaldi, Rajiv Sethi, Gil Skillman, Peter Skott, Tsuyoshi Tsuru, and audiences at LSE, UCL, Rome (Tor Vergata), Liverpool, UMass Amherst, UMass Boston, Siena, Leeds, Toyama, Udine, APE workshop (QMUL), WAPE conference (Amherst), INET conference (Waterloo), AEA conference (Chicago), G-COE Conference on Equality and Welfare (Hitotsubashi), HM conference (New York), EEA conference (New York), FMM conference (Berlin), Quantitative Methods workshop (QMUL), AHE conference (Bristol), and EAGT conference (Korea) for comments and suggestions. Grau worked on this project during his tenure as a professor in the Department of Economics at the Universidad de Chile before joining the Chilean Government as Minister. The views expressed here, therefore, are those of the authors and do not represent those of the Ministry of Economy, Development, and Tourism, or the Chilean government. The usual disclaimer applies.

[†]Department of Political Economy, King's College London (daniele.girardi@kcl.ac.uk)

[‡]Ministry of Economy, Development, and Tourism, Government of Chile (ngrau@fen.uchile.cl)

[§]School of Economics and Finance, Queen Mary University of London (r.veneziani@qmul.ac.uk)

[¶]Department of Economics, University of Massachusetts Amherst; The Institute of Economic Research, Hitotsubashi University, Graduate School of Economics and Management, Tohoku University, and School of Management, Kochi University of Technology (nyoshihara@econs.umass.edu)

"Perhaps it is in the nature of workers to translate themselves into what they work upon" (Salman Rushdie, *Two years, eight months & twenty nights*, London, Random House, p.39.)

1 Introduction

The notion of exploitation is prominent in some of the social sciences and in political discourse. It is central in a number of debates, ranging from analyses of labour relations, especially focusing on the weakest segments of the working population (ACLU 2023; ILO 2024); to controversies on drug-testing and on the price of life-saving drugs, especially in developing countries (Hawkins and Emanuel 2008; Lamkin and Elliott 2018); to ethical issues arising in surrogate motherhood (Tan 2020; Howard 2023; Lee 2023). Economic exploitation involving unfair compensation of work and/or excessive working hours is sometimes codified as a criminal offence (see, for example, Art.603-bis of the Italian Criminal Code).

The concept of exploitation is also central in progressive politics. The manifestos of both the UK Labour Party and the German SPD, for example, advocate the end of exploitation in labour relations (SPD 2021; Labour Party 2024), as does the document approved by the Spanish PSOE during its last congress (PSOE 2021). The fight against economic exploitation is repeatedly indicated as a priority for three of the main parties of the European Left.

Yet, whereas economic exploitation is extensively discussed in philosophy; political science; and sociology,[1] it is rather marginal in economics where it is taken to denote, at best, fraudulent, coercive, or otherwise anticompetitive practices. Beyond these problematic but somewhat peripheral instances, the concept of exploitation is widely considered to be obscure and metaphysical; lacking any clear normative content; and often based on logically inconsistent foundations, as in the Marxian approach.[2]

This paper proves that, contrary to the received view, the concept of exploitation as the unequal exchange of labour can be given a precise and consistent theoretical definition which is uniquely characterised based on a small set of normatively intutitive properties. From this

[1] The literature in each discipline is too vast for a comprehensive set of references. An illustrative, but far from comprehensive selection of *recent* contributions includes: in philosophy, Vrousalis (2013, 2022); Horton (2019); Ferguson (2021) – see Zwolinski, Ferguson, and Wertheimer (2022) for a survey; in political science, Gourevitch (2018); Valdez (2020); Gerver (2022); Bryan (2023); Temin (2024); Chan (2024); in sociology, Bartley and Child (2014); Mears (2015); Tomaskovic-Devey, Hällsten, and Avent-Holt (2015); Wodtke (2016); Desmond and Wilmers (2019); Sauer, Valet, Shams, and Tomaskovic-Devey (2021).

[2] A notable recent exception is McGee (2025), which analyses racial relations through the lens of exploitation theory. The definition of exploitation implicitly adopted by McGee (2025) is consistent with the one developed here.

definition, a well-defined empirical measure of exploitation intensity at the individual level can be derived. The exploitation intensity index provides novel normative insights on actual economies and may contribute to explain socioeconomic outcomes of interest.

We start by providing the first axiomatic characterisation of a definition of exploitation as the unequal exchange of labour in the context of general equilibrium models in the tradition of Morishima (1974), Flaschel (1983), and, especially, Roemer (1980, 1982). Formally, a definition of exploitation is a mapping which – in the general equilibrium of any convex economy – identifies the set of exploiters and the set of exploited agents (and those who belong to neither category).[3] A domain condition named *Labour Exploitation* identifies the set of definitions according to which exploitative relations are characterised by systematic differences between the labour that agents contribute to the economy and the labour 'received' by them. The former is given by the amount of (effective) labour performed in production. The latter is given by the amount of labour contained, or embodied, in what may be called *exploitation reference bundles* (henceforth, ERBs).[4]

Within the domain of definitions of labour exploitation, Theorem 1 proves that only one definition satisfies three theoretically robust, formally weak, and normatively appealing properties. *Scale Invariance* and *Independence* are conceptually analogous to similar axioms in social choice theory and inequality measurement. They constrain the way in which the ERBs, and the labour they contain, change across equilibria. *Scale Invariance* says that, for given equilibrium prices, technology, aggregate endowments, and aggregate production activity, proportional changes in individual endowments and economic activities should yield equiproportional changes in the amount of labour received by agents. *Independence* states that, for given production set, set of agents, aggregate endowments, and price vector, if either (i) preferences change, but consumption choices remain the same; or (ii) consumption choices change but all other choices *and* preferences remain the same; then the labour received by agents in the ERBs should be unchanged. Finally, *Relational Exploitation* simply says that whenever someone is exploited, there must be an exploiter, and vice versa.

Theorem 1 proves that the only definition consistent with all three axioms identifies an agent as exploited (an exploiter) if and only if their share of total labor contribution exceeds (is less than) their share of total income. This definition corresponds to the so-

[3]We note in passing that we focus exclusively on the *distributive* dimension of exploitation. This is merely a choice of analytical focus as power asymmetries are arguably an essential ingredient of exploitative relations. For a thorough discussion, see Veneziani (2007); Vrousalis (2013, 2022).

[4]Conceptually, the ERBs play a similar role in exploitation theory as the *egalitarian reference bundles* in the theory of fair allocation (Pazner and Schmeidler 1978; Moulin 1987; Thomson 1994; Fleurbaey 1995; Fleurbaey and Maniquet 1999; Maniquet and Sprumont 2004).

called 'New Interpretation' (Duménil 1980, 1984; Foley 1982, 1986). It embodies a notion of proportionality between contribution and reward that has been explored in normative economics (Roemer and Silvestre 1993; Moulin 1990) and whose philosophical roots can be traced back to Aristotle (Maniquet 2002) and Kant (Roemer 2019).[5]

From Theorem 1, it is possible to derive an index that measures *exploitation intensity* at the individual level, based on the distance between the proportion of labour contributed by an agent relative to total labour, and the proportion of income she receives. Importantly, the resulting exploitation intensity index is based on variables that can be estimated empirically.

We then use nationally representative survey data to provide the first empirical analysis of exploitation intensity at the individual level, focusing on the United States during 1975-2022. First, the data suggest that most people are exploited: the fraction of the population with a positive exploitation index hovers slightly above 75%, with no clear trend, over the period. The exploitation intensity index for the median American is around 0.50 in 2022: the fraction of labour contributed by the median American is approximately 50 percentage points above the proportion of income received.

Moreover, the 1975-2022 period has seen a marked increase in the gap between the top and bottom percentiles of exploitation intensity. The largest increase in exploitation intensity has been experienced by the most exploited percentile and by those in the middle of the distribution, while the better off have further improved their position. According to our estimates, exploitation intensity increased by almost 30 percent for the median American, almost 15 percent for the 90th percentile and over 33 percent for the 99th percentile; but it decreased by over 20 percent per the least exploited 10% of the population.

An analysis based on the Gini index suggests that inequality in exploitation increased more than income inequality over the 1975-2022 period. The Gini index for exploitation intensity increased by around 22 percent over this period, while in the same survey data the Gini index for total income increased by 1.6 percent in the whole sample and by 19 percent among full-time workers.

In order to investigate the determinants of exploitation, we regress exploitation intensity on a set of binary variables capturing potentially relevant individual characteristics, controlling for a full set of age-by-year fixed effects.

We find that women and blacks tend to be substantially more exploited. The gender

[5]Perhaps more surprisingly, it is possible to prove that all of the main definitions of labour exploitation in the literature – including the classic approaches proposed by Morishima (1974); Roemer (1981, 1982); Flaschel (1983) – satisfy *Scale Invariance* and *Independence* (see Appendix B). Therefore, it is the rather mild and uncontroversial axiom *Relational Exploitation* that rules out all alternative approaches except the 'New Interpretation'.

exploitation gap has decreased significantly over the sample period but it still stands at 18 percent in 2005-2022. The racial exploitation gap is around 10 percent over the whole period, with no clear trend. Perhaps unsurprisingly, rentiers (defined as individuals who receive positive capital income) and entrepreneurs and the retired tend to be substantially less exploited than individuals whose income only consists of wages.

We then present a descriptive analysis suggesting that exploitation intensity does matter, in the sense of carrying predictive power for important socio-economic outcomes at the individual level. We estimate a number of regressions of a variety of socio-economic outcomes on percentile of exploitation intensity and a full set of age and year fixed effects, with and without controlling for percentile of income. Results show that more exploited individuals report lower happiness, worse health, lower job satisfaction and are more likely to identify as working class. They are more likely to support the Democratic party in the 1974-1995 period but not after 1995. Importantly, the correlation between exploitation and these outcomes is robust to controlling for income (with the sole exception of job satisfaction).[6]

Our main analysis uses hours worked to measure labour contribution, yielding an intuitively appealing index that is easy to interpret and measure. However, we also consider two alternative measures of labour contribution that attempt to account for differences in effort and ability. Taken together, the three approaches we consider are likely to provide a reasonably accurate and comprehensive picture of the empirical dynamics of exploitation, as defined in our theoretical framework, and reassuringly our key results are robust to using any of the three measures.

The exploitation intensity index that we derive and estimate empirically represents a meaningful new measure of well-being and inequality, complementary to existing ones and able to jointly take into account the distributions of income and work time. A measure of material well-being based on work hours and total income is intuitively appealing, given the importance of these two variables for human welfare. It can provide additional insights relative to traditional measures based on real income or hourly wages: unlike the first, it takes into account hours worked, and unlike the second, it accounts for unearned incomes. Conceptually, income inequality indexes measure deviations from a benchmark of equal income distribution, whereas exploitation intensity measures deviations from a benchmark of income proportional to individual contribution. Seen in this light, the contribution of this paper is to provide explicit theoretical foundations for such a measure, and uncover its formal connection to the concept of exploitation as the unequal exchange of labour.

[6]The size of these correlations is non-negligible, and roughly comparable to the size of the correlation between these variables and income.

For these reasons, this paper contributes both to the theoretical literature on defining exploitation,[7] *and* more broadly to a vast literature on measuring material well-being and inequality, and on their recent empirical trends.[8] Our empirical findings are largely in line with the literature on the recent evolution of income inequality in the US (documented for example in Hoffmann et al. 2020 using the same survey data we employ here). Like income inequality, inequality in our exploitation index increased in the USA since the 1980s. However, inequality in exploitation intensity increased *more* than income inequality: our proposed index shows that the recent increase in inequality in the US is larger when considering total income received in relation to labour provided. Importantly, this finding is largely robust to the alternative ways of measuring labour contribution discussed above.

The rest of the paper is structured as follows. Section 2 lays out the formal framework. Section 3 formalises the concept of definition of exploitation and presents the main domain condition identifying definitions of labour exploitation. Section 4 presents the main characterisation. Section 5 defines the exploitation intensity index and describes our empirical strategy. Section 6 derives the distribution of exploitation in the US during 1975-2022. Section 7 investigates the relation between exploitation intensity and a number of socioeconomic outcomes of interest. Section 8 concludes.

2 The model

This section presents a generalisation of Roemer's (1980; 1982) classic economies and of the related equilibrium notion.

2.1 Production

Let $\mathbb{R}$ ($\mathbb{R}_+$) be the set of (nonnegative) real numbers. Let $\mathbf{0}$ denote the null vector.[9] Production technology is freely available to all agents, who can operate any activity in the production set $\mathcal{P}$, which has elements, *activities*, of the form $\alpha = (-\alpha_l, -\underline{\alpha}, \overline{\alpha})$ where $\alpha_l \in \mathbb{R}_+$ is the *effective* labour input; $\underline{\alpha} \in \mathbb{R}^n_+$ are the inputs of the produced goods; and $\overline{\alpha} \in \mathbb{R}^n_+$ are the outputs of the n goods. Production displays constant returns to scale: $\mathcal{P}$ is a closed convex

[7] For example, Morishima (1974); Roemer (1980, 1982); Duménil (1980, 1984); Foley (1982, 1983); Flaschel (1983); Fleurbaey (2014); Veneziani (2007); Yoshihara (2010); Galanis, Veneziani, and Yoshihara (2019)

[8] For example, Kuznets (1954); Atkinson (1970); Piketty and Saez (2003); Stiglitz, Sen, and Fitoussi (2009); Milanovic (2016); Alvaredo, Chancel, Piketty, Saez, and Zucman (2017); Piketty, Saez, and Zucman (2018); Hoffmann, Lee, and Lemieux (2020).

[9] All vectors are columns, unless otherwise specified.

cone in $\mathbb{R}^{2n+1}$ with $\mathbf{0} \in \mathcal{P}$.[10]

Let $\partial\mathcal{P} \equiv \{\alpha \in \mathcal{P} \mid \nexists \alpha' \in \mathcal{P} \text{ s.t. } \alpha' > \alpha\}$ denote the frontier of $\mathcal{P}$,[11] and let the net output vector arising from α be denoted as $\widehat{\alpha} \equiv \overline{\alpha} - \underline{\alpha}$. For any $c \in \mathbb{R}^n_+$, the set of activities that produce at least c as net output is:

$$\phi(c) \equiv \{\alpha \in \mathcal{P} \mid \widehat{\alpha} \geqq c\}.$$

2.2 Agents

The economy comprises a set of agents $\mathcal{N} = \{1, ..., N\}$ where N is assumed to be sufficiently large.[12] Agents are endowed with unequal amounts of physical and human capital, and they produce, consume, and trade labour. On the production side, they can either sell their labour-power or hire workers to work on their capital, or they can be self-employed and work on their own assets. Formally, for all $\nu \in \mathcal{N}$, let $s^\nu > 0$ be agent ν's skill level and let $\omega^\nu \in \mathbb{R}^n_+$ be the vector of productive assets inherited by ν. Then, $\alpha^\nu = (-\alpha^\nu_l, -\underline{\alpha}^\nu, \overline{\alpha}^\nu) \in \mathcal{P}$ is the activity operated by ν as a self-employed producer, where $\alpha^\nu_l = s^\nu a^\nu_l$ and a^ν_l is the labour *time* expended by ν; $\beta^\nu = \left(-\beta^\nu_l, -\underline{\beta}^\nu, \overline{\beta}^\nu\right) \in \mathcal{P}$ is the activity that ν operates by hiring (effective) labour β^ν_l; $\gamma^\nu = s^\nu l^\nu$ is ν's (effective) labour supply, where l^ν is the labour *time* supplied by ν on the market. Thus, $\lambda^\nu = (a^\nu_l + l^\nu)$ is the total amount of labour time expended by ν, and $\Lambda^\nu = \alpha^\nu_l + \gamma^\nu = s^\nu \lambda^\nu$ is the total amount of effective labour performed by ν, either as a self-employed producer or working for some other agent.[13]

On the consumption side, let $c^\nu \in \mathbb{R}^n_+$ be agent ν's vector of consumption goods. Total labour hours expended by each agent cannot exceed the total amount of time available, which is normalised to one. Agent ν's welfare is representable by a function $u^\nu : \mathbb{R}^n_+ \times [0,1] \to \mathbb{R}_+$, which is increasing in consumption and decreasing in labour time, continuous, and quasi-concave. For the sake of simplicity, and with no loss of generality, u^ν is assumed to be strictly monotonic in at least one of the first n arguments, for all ν.[14]

[10] A formal exposition and discussion of the properties of $\mathcal{P}$ is in Appendix A.

[11] For all vectors $x, y \in \mathbb{R}^q$, $x \geqq y$ if and only if $x_i \geqq y_i$ $(i = 1, \ldots, q)$; $x \geq y$ if and only if $x \geqq y$ and $x \neq y$; $x > y$ if and only if $x_i > y_i$ $(i = 1, \ldots, q)$.

[12] This assumption is conceptually analogous to Roemer's (1982) *Assumption of a Large Economy*: it rules out some very special cases and it is without loss of generality both theoretically and formally. Theoretically, we are interested in empirical measures capturing exploitation in actual economies. Formally, the main characterisation result holds for any $N \geqq 2$ provided a standard *Replication Invariance* axiom is imposed (see Appendix D).

[13] The model does not include different types of labour to be used in production. This is only for simplicity: this additional source of heterogeneity can be dealt with, albeit at the cost of a substantial increase in technicalities (see Yoshihara and Veneziani (2023)).

[14] For a characterisation in the special case of economies in which agents simply minimise labour, see

Let p denote the $1 \times n$ vector of commodity prices and let w denote the wage rate *per unit of effective labour*. Given (p, w), each agent ν chooses a plan $\xi^\nu \equiv (\alpha^\nu, \beta^\nu, \gamma^\nu, c^\nu)$ to maximise her welfare subject to the constraint that (1) net income is sufficient for consumption plans; (2) wealth is sufficient to purchase the inputs necessary for production plans; (3) production plans are technically feasible; and (4) consumption and leisure are feasible. Formally, each agent $\nu \in \mathcal{N}$ solves:[15]

$$MP^\nu : \max_{\xi^\nu = (\alpha^\nu, \beta^\nu, \gamma^\nu, c^\nu)} u^\nu (c^\nu, \lambda^\nu)$$

subject to

$$\begin{aligned} [p(\overline{\alpha}^\nu - \underline{\alpha}^\nu)] + \left[p\left(\overline{\beta}^\nu - \underline{\beta}^\nu\right) - w\beta_l^\nu\right] + [w\gamma^\nu] &= pc^\nu, & (1)\\ p\left(\underline{\alpha}^\nu + \underline{\beta}^\nu\right) &\leqq p\omega^\nu, & (2)\\ \alpha^\nu, \beta^\nu &\in \mathcal{P}, & (3)\\ c^\nu \in \mathbb{R}^n_+,\ \lambda^\nu &\in [0,1]. & (4) \end{aligned}$$

MP^ν is a generalisation of similar optimisation programmes in Roemer (1980, 1982). It incorporates a standard view of individual behaviour but it differs from traditional models in two respects. First, MP^ν incorporates the simultaneous role of economic actors as consumers (see, in particular, (1) and (4)) and producers (see, in particular, (2) and (3)), so that no separate consideration of firms is necessary.[16] Second, it explicitly takes into account the time structure of the production process. It is thus assumed that, at the beginning of the period, agents need to lay out in advance the capital needed for production and can do so only by using their own wealth (see (2)).[17] Production then takes place and gross revenues (including wages and profits) can be used to finance consumption and the reproduction of initial wealth at the end of the period (see (1)).[18]

2.3 Equilibrium

Let $E\langle \mathcal{P}, \mathcal{N}, (u^\nu, s^\nu, \omega^\nu)_{\nu \in \mathcal{N}} \rangle$, or as a shorthand notation E, denote the economy with technology $\mathcal{P}$, agents $\mathcal{N}$, utility functions $(u^\nu)_{\nu \in \mathcal{N}}$, labour skills $(s^\nu)_{\nu \in \mathcal{N}}$, and physical endowments

Yoshihara and Veneziani (2009).

[15]The first constraint is written as an equality given the assumptions on the monotonicity of u^ν.

[16]As shown in Lemma 1 below, profit maximisation is a corollary of MP^ν.

[17]A credit market may be introduced but it would not change the main results and the structure of exploitative relations. See Roemer (1981, 1982).

[18]Because of the time structure of production, prices may differ at the beginning and at the end of the period. Given the focus of this paper, however, it is appropriate to analyse stationary equilibria, and assume agents rationally to expect prices to be constant, as in Roemer (1980, 1981, 1982).

$(\omega^\nu)_{\nu\in\mathcal{N}}$. Let the set of all such economies be denoted by $\mathcal{E}$. Following Roemer (1980, 1981, 1982), the equilibrium concept can be defined.

Definition 1. A *reproducible solution* (RS) for $E \in \mathcal{E}$ is a price vector (p, w) and an associated profile of actions $(\xi^\nu)_{\nu\in\mathcal{N}} = (\alpha^\nu, \beta^\nu, \gamma^\nu, c^\nu)_{\nu\in\mathcal{N}}$ such that:

(i) ξ^ν solves MP^ν for all ν (optimality);

(ii) $\sum_{\nu\in\mathcal{N}} \left(\underline{\alpha}^\nu + \underline{\beta}^\nu\right) \leqq \sum_{\nu\in\mathcal{N}} \omega^\nu$ (feasibility);

(iii) $\sum_{\nu\in\mathcal{N}} \beta_l^\nu = \sum_{\nu\in\mathcal{N}} \gamma^\nu$ (labour market equilibrium);

(iv) $\sum_{\nu\in\mathcal{N}} \left(\widehat{\alpha}^\nu + \widehat{\beta}^\nu\right) \geqq \sum_{\nu\in\mathcal{N}} c^\nu \geq \mathbf{0}$ (reproducibility).

At a RS, (i) every agent optimises; (ii) there are enough resources for aggregate production plans; and (iii) the labour market clears. Condition (iv) states that aggregate net outputs should at least suffice for aggregate consumption without depleting physical endowments.[19] Indeed, although a RS is defined as a temporary equilibrium in a static general equilibrium framework, it can be seen as a one-shot slice of a *stationary equilibrium* in a dynamic general equilibrium framework.[20]

By the assumptions on u^ν, it immediately follows that both the wage and the prices of all goods must be nonnegative, and at least one good must have a strictly positive price. Further, let $\pi = \max_{\alpha\in\mathcal{P}} \frac{p\widehat{\alpha} - w\alpha_l}{p\underline{\alpha}}$ denote the maximum profit rate that can be obtained at prices (p, w). Lemma 1 derives some useful properties of the equilibria of the economy.

Lemma 1. *Let* $\left((p, w), (\xi^\nu)_{\nu\in\mathcal{N}}\right)$ *be a RS for* $E \in \mathcal{E}$ *such that* $\sum_{\nu\in\mathcal{N}} c^\nu \geq \mathbf{0}$. *Then, (i)* $p\widehat{\alpha} - w\alpha_l \geqq 0$ *for some* $\alpha \in \mathcal{P}\setminus\{\mathbf{0}\}$, *and (ii)* $\pi = \frac{p\widehat{\alpha}^\nu - w\alpha_l^\nu}{p\underline{\alpha}^\nu} = \frac{p\widehat{\beta}^\nu - w\beta_l^\nu}{p\underline{\beta}^\nu}$ *for all* $\nu \in \mathcal{N}$.

The proof of Lemma 1 is straightforward and therefore omitted. Intuitively, by individual optimality, in equilibrium agents will not operate any activities that yield negative profits and only choose activities that yield the maximum profit rate.

[19] Condition (iv) is equivalent to requiring that the vector of social endowments does not decrease component-wise, because it is equivalent to $\sum_{\nu\in\mathcal{N}} \left[\omega^\nu - \left(\underline{\alpha}^\nu + \underline{\beta}^\nu\right) + \left(\overline{\alpha}^\nu + \overline{\beta}^\nu - c^\nu\right)\right] \geqq \sum_{\nu\in\mathcal{N}} \omega^\nu$, which states that aggregate stocks at the beginning of next period should not be smaller than aggregate stocks at the beginning of the current period.

[20] For different dynamic generalisations of the concept of RS, see Roemer (1980); Fleurbaey (1996); Veneziani (2007); Veneziani and Yoshihara (2017); Galanis et al. (2019).

3 Defining labour exploitation

A definition of exploitation is a rule that, for any given economy, and any given equilibrium allocation of this economy, identifies the exploitation status of all agents. Formally, for any set of agents $\mathcal{N}$, let $\mathcal{T}_{\mathcal{N}}$ be the set of all conceivable partitions of $\mathcal{N}$. Let $\aleph$ be the universal set comprising all conceivable sets of agents and let $\mathcal{T} = \cup_{\aleph}\mathcal{T}_{\mathcal{N}}$. For each $E \in \mathcal{E}$, let $\mathcal{RS}_E$ be the set of reproducible solutions of E and let $\mathcal{RS} = \cup_{\mathcal{E}}\mathcal{RS}_E$.

Definition 2. A *definition of exploitation* is a mapping $d : \mathcal{E} \times \mathcal{RS} \longrightarrow \mathcal{T}$ such that for each $E = E\langle \mathcal{P}, \mathcal{N}, (u^\nu, s^\nu, \omega^\nu)_{\nu\in\mathcal{N}}\rangle \in \mathcal{E}$ and $\left((p, w), (\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, $d(E, \left((p, w), (\xi^\nu)_{\nu\in\mathcal{N}}\right)) = \{\mathcal{N}^{ter}, \mathcal{N}^{ted}, \mathcal{N}^n\} \in \mathcal{T}_{\mathcal{N}}$ where $\mathcal{N}^{ter}$ is the set of exploiters, $\mathcal{N}^{ted}$ is the set of exploited agents, and $\mathcal{N}^n$ is the set of agents who are neither exploiters nor exploited.

Let $\mathcal{D}$ denote the set of conceivable definitions of exploitation. As Definition 2 is very general, and it allows one to formalise all views on exploitation, $\mathcal{D}$ is very large. In principle, there are many ways of identifying the subsets $\{\mathcal{N}^{ter}, \mathcal{N}^{ted}, \mathcal{N}^n\}$ at any given equilibrium allocation. For example, a libertarian may insist that mutually beneficial transactions in perfectly competitive markets are non-exploitative and endorse the constant mapping $d(E, \left((p, w), (\xi^\nu)_{\nu\in\mathcal{N}}\right)) = \{\emptyset, \emptyset, \mathcal{N}\}$ for all $E \in \mathcal{E}$ and $\left((p, w), (\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$.

In this paper, a novel, general axiomatic framework is developed in order to analyse exploitation theory. The adoption of an axiomatic method allows us to adjudicate alternative approaches by starting from first principles, thus explicitly discussing the intuitions underlying different definitions $d \in \mathcal{D}$.

To be specific, we are interested in a subset of $\mathcal{D}$, namely those approaches that focus on the *unequal exchange of labour* (henceforth, UE), including the classic definitions by Morishima (1974); Roemer (1980, 1982); Duménil (1980, 1984); Foley (1982, 1983), and Flaschel (1983, 2010), among the others. While these definitions are very different, all labour-based approaches share a common conceptual structure: exploitative relations are characterised by systematic differences between the labour that agent ν contributes to the economy, Λ^ν, and the labour 'received' by them, which is given by the amount of labour contained, or embodied, in some relevant consumption bundle(s). Therefore, in order to define exploitation status, it is necessary both to select the relevant reference bundle(s) and to identify their labour content. For example, one may argue that exploitation theory should focus on the bundles *actually* purchased by agents, c^ν, and measure the labour contained in such bundles by using the Leontief employment multipliers, ψ derived from the production technique(s) *actually* used. Then, an agent ν is exploited if and only if $\Lambda^\nu > \psi c^\nu$ and an exploiter if and only if $\Lambda^\nu < \psi c^\nu$.

The next property sets some weak restrictions *both* on the choice of the reference bundle(s) *and* on the definition of their labour content and it defines the domain of labour-based definitions of exploitation.[21]

For any $p, c \in \mathbb{R}^n_+$, let $\mathcal{B}(p,c) \equiv \left\{x \in \mathbb{R}^n_+ \mid px = pc\right\}$ be the set of bundles that cost exactly as much as c at prices p. Then:

Labour Exploitation (LE): *For every $E \in \mathcal{E}$ and $\left((p,w), (\xi^\nu)_{\nu \in \mathcal{N}}\right) \in \mathcal{RS}_E$, there exists a profile $(\overline{c}^\nu, \underline{c}^\nu)_{\nu \in \mathcal{N}}$ such that for each $\nu \in \mathcal{N}$, $\overline{c}^\nu, \underline{c}^\nu \in \mathcal{B}(p, c^\nu)$ and there exist $\alpha^{\overline{c}^\nu} \in \phi(\overline{c}^\nu) \cap \partial\mathcal{P}$ and $\alpha^{\underline{c}^\nu} \in \phi(\underline{c}^\nu) \cap \partial\mathcal{P}$ such that $\alpha_l^{\underline{c}^\nu} \geqq \alpha_l^{\overline{c}^\nu}$ and*

$$\nu \in \mathcal{N}^{ter} \Leftrightarrow \Lambda^\nu < \alpha_l^{\overline{c}^\nu};$$
$$\nu \in \mathcal{N}^{ted} \Leftrightarrow \Lambda^\nu > \alpha_l^{\underline{c}^\nu}.$$

Furthermore, if $(u^\nu, \omega^\nu, s^\nu, \xi^\nu) = (u^\mu, \omega^\mu, s^\mu, \xi^\mu)$ for all $\nu, \mu \in \mathcal{N}$, then $\left(\overline{c}^\nu, \underline{c}^\nu, \alpha_l^{\overline{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right) = \left(\overline{c}^\mu, \underline{c}^\mu, \alpha_l^{\overline{c}^\mu}, \alpha_l^{\underline{c}^\mu}\right)$ for all $\nu, \mu \in \mathcal{N}$.

LE requires that, at any equilibrium, the sets $\mathcal{N}^{ter}$ and $\mathcal{N}^{ted}$ be determined by identifying two profiles of (possibly identical) nonnegative vectors $\overline{c}^\nu, \underline{c}^\nu \in \mathbb{R}^n_+$, that may be called the *exploitation reference bundles* (hereafter, ERBs).[22]

The ERBs must be *just affordable* at prices p, w by optimising agents ($\overline{c}^\nu, \underline{c}^\nu \in \mathcal{B}(p, c^\nu)$), which captures the idea that the amount of labour that each agent receives depends on their income. The ERBs may – but need not – correspond to the bundle actually chosen by each agent: **LE** is much weaker in that it allows for more than one reference bundle *for every agent* and it only requires that the ERBs be *potentially* affordable.

The ERBs must also be *technically feasible* ($\alpha^{\overline{c}^\nu} \in \phi(\overline{c}^\nu) \cap \partial\mathcal{P}, \alpha^{\underline{c}^\nu} \in \phi(\underline{c}^\nu) \cap \partial\mathcal{P}$), and the amount of labour received by agents is given by the labour necessary to produce them as net output, $\left[\alpha_l^{\overline{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right]$. This amount must be uniquely determined, but (i) it is not specified and there may be many ways of producing $\overline{c}^\nu, \underline{c}^\nu$; (ii) it need not be a scalar as the interval $\left[\alpha_l^{\overline{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right]$ may be non-degenerate; and (iii) it may be derived based either on the production technique(s) actually used, or on some counterfactual activity.[23]

Then, agent ν is an *exploiter* if and only if ν contributes less than the minimum amount

[21] For a thorough discussion of the philosophical foundations of **LE**, see Veneziani and Yoshihara (2018).

[22] The set $\mathcal{N}^n$ is determined as a residual and by **LE** it may be of non-zero measure if $\alpha_l^{\underline{c}^\nu} > \alpha_l^{\overline{c}^\nu}$.

[23] Once the ERBs, $\overline{c}^\nu, \underline{c}^\nu$, are identified, the existence of $\alpha^{\overline{c}^\nu}, \alpha^{\underline{c}^\nu}$ is guaranteed by the assumptions on $\mathcal{P}$ (see **A2** and **A3** in Appendix A).

of labour that ν can receive via her 'net income', $\alpha_l^{\bar{c}^\nu}$; whereas agent ν is *exploited* if and only if ν contributes more than the maximum amount of labour that ν can receive via her 'net income', $\alpha_l^{c_t^\nu}$.

Finally, observe that **LE** states that if all agents are identical and they choose exactly the same actions, then the ERBs and associated labour contents are the same. This restriction can be seen as conceptually similar to standard anonymity properties in normative economics since it requires that agents' identities be irrelevant in identifying the criteria to evaluate their exploitation status (while being silent on their actual exploitation status).

LE represents an appropriate condition to identify the domain of labour-based definitions of exploitation: it is formally weak and, as shown in Appendix B, it captures the key insights of exploitation theory that are shared by *all* of the main labour-based approaches.

If $\mathcal{D}$ is the set of conceivable definitions of exploitation, **LE** is a restriction on $\mathcal{D}$ that identifies definitions of *labour exploitation*.[24]

Definition 3. A definition of exploitation d that satisfies **LE** is a *definition of labour exploitation*. $\mathcal{D}_L \subset \mathcal{D}$ is the set of definitions of exploitation that satisfy **LE**.

In the next section, we shall identify some desirable properties of definitions of labour exploitation – that is, restrictions on the mapping $d \in \mathcal{D}_L$.

4 An axiomatic approach

The first axiom captures an arguably essential feature of *any* theory of exploitation.

Relational Exploitation (RE): For every $E \in \mathcal{E}$ and $\left((p,w), (\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, $\mathcal{N}^{ter} \neq \varnothing$ *if and only if* $\mathcal{N}^{ted} \neq \varnothing$.

Formally, axiom **RE** imposes a rather weak restriction on $\mathcal{D}_L$. Theoretically, it captures the crucial relational aspect inherent in exploitative relations, such that if an agent is exploited, she must be exploited *by someone*, and vice versa any exploiters must be exploiting *someone*.

In order to introduce the next properties, let $\mathcal{E}(\mathcal{P}^*; \mathcal{N}^*; s; \omega) \subset \mathcal{E}$ be the set of economies with the same production set $\mathcal{P}^*$, set of agents $\mathcal{N}^*$, and aggregate endowments of labour, s, and commodity inputs ω. Thus, if $E, E' \in \mathcal{E}(\mathcal{P}^*; \mathcal{N}^*; s; \omega)$ then $\mathcal{P} = \mathcal{P}' = \mathcal{P}^*$, $\mathcal{N} = \mathcal{N}' = \mathcal{N}^*$, $\sum_{\nu\in\mathcal{N}} \omega^\nu = \sum_{\nu\in\mathcal{N}'} \omega'^\nu = \omega$, and $\sum_{\nu\in\mathcal{N}} s^\nu = \sum_{\nu\in\mathcal{N}'} s'^\nu = s$.

[24]Libertarian approaches defining exploitation based on prior rights violations (Steiner 1984), approaches focusing on public good contributions (Fleurbaey 2014), or Roemer's (1982) property-relations definition, do not satisfy **LE** and thus do not belong to $\mathcal{D}_L$.

For any $E \in \mathcal{E}$ and any $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, let $\alpha^{p,w} + \beta^{p,w} \equiv \sum_{\nu\in\mathcal{N}} (\alpha^\nu + \beta^\nu)$ denote the aggregate equilibrium production activity.

The second axiom is analogous to standard scale invariance conditions in social choice.

Scale Invariance (SI): *Consider any* $E, E' \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$ *and any* $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$ *and* $\left((p,w),(\xi'^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E'}$ *such that* $\alpha^{p,w} + \beta^{p,w} = \alpha'^{p,w} + \beta'^{p,w}$ *and for all* $\nu \in \mathcal{N}$, *there exists* $\chi^\nu > 0$ *such that* $s^\nu = \chi^\nu s'^\nu$, $\omega^\nu = \chi^\nu \omega'^\nu$, and $\xi^\nu = \chi^\nu \xi'^\nu$. *Let* $(\overline{c}^\nu, \underline{c}^\nu)_{\nu\in\mathcal{N}}$ *and* $(\overline{c}'^\nu, \underline{c}'^\nu)_{\nu\in\mathcal{N}}$ *be the corresponding ERBs. Then,* $\left(\alpha_l^{\overline{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right) = \chi^\nu \left(\alpha_l^{\overline{c}'^\nu}, \alpha_l^{\underline{c}'^\nu}\right)$ *for all* $\nu \in \mathcal{N}$.

SI says that, for given equilibrium prices, technology, aggregate endowments, and aggregate production activity, proportional changes in individual endowments and economic activities should yield equiproportional changes in the amount of labour received by agents. Given **LE**, it is immediate to see that this implies that the sets of exploiters and exploited agents are invariant with respect to equiproportional changes in individual choices and endowments.

The final axiom is conceptually analogous to standard independence conditions.

Independence (IND): *Consider any* $E, E' \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$ *and any* $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$ *and* $\left((p,w),(\xi'^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E'}$ *such that* $\alpha^{p,w} + \beta^{p,w} = \alpha'^{p,w} + \beta'^{p,w}$. *Let* $(\overline{c}^\nu, \underline{c}^\nu)_{\nu\in N}$ *and* $(\overline{c}'^\nu, \underline{c}'^\nu)_{\nu\in\mathcal{N}}$ *be the corresponding ERBs. If either (i)* $E \neq E'$ *and* $c^\nu = c'^\nu$ *for all* $\nu \in \mathcal{N}$ *, or (ii)* $E = E'$ *and* $(\Lambda^\nu, \alpha^\nu, \beta^\nu, \gamma^\nu) = (\Lambda'^\nu, \alpha'^\nu, \beta'^\nu, \gamma'^\nu)$ *for all* $\nu \in \mathcal{N}$, *then* $\left(\alpha_l^{\overline{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right) = \left(\alpha_l^{\overline{c}'^\nu}, \alpha_l^{\underline{c}'^\nu}\right)$ *for all* $\nu \in \mathcal{N}$.

In order to interpret **IND**, note that it does *not* constrain agents' exploitation status. Rather it imposes some mild conditions on the way in which such status is determined – namely, on the ERBs and the associated labour content. According to **LE**, the latter depend in general on the characteristics of the economy and on the actual equilibrium considered, and should capture both technical conditions of production and the agents' feasible choices. **IND** considers economies with the same set of agents and the same aggregate endowments, as well as identical production sets, and equilibria with the same equilibrium price vector and aggregate production activity.

Under these rather stringent conditions, part (i) says that if the consumption vectors of all agents are also identical in the equilibria of the two economies, then the ERBs should be such that their associated amounts of labour are identical. In other words, if two economies share the same economic structure, except for arguably irrelevant factors (the distributions of endowments of productive inputs and the profile of utility functions), and have the same equilibrium prices and aggregate equilibrium activities, then the upper and lower bounds of the labour received should be identical, provided the consumption bundles chosen by all agents are also identical.

Part (ii) relaxes the latter restriction and allows the agents' consumption bundles to vary across equilibria but it now requires the two economies and all other optimal choices to be identical. In other words, we require the upper and lower bounds of the labour received to be independent of consumption choices *only* in the very special case in which, at a given equilibrium price vector, agents have multiple (indeed, generically a continuum of) optimal consumption bundles. Part (ii) says that the definition of exploitation status should not depend on the possibly arbitrary selection of one consumption bundle from a (non-single valued) optimal correspondence – should this actually be possible in equilibrium.

In other words, **IND** specifies the conditions under which preferences and consumption choices matter in the determination of the criteria to identify exploitation status. For given production set, set of agents, aggregate endowments, and price vector, if either (i) preferences change, but consumption choices remain the same; or (ii) consumption choices change but all other choices *and* preferences remain the same; then the ERBs should be such that the upper and lower bounds of the labour received by each agent should be unchanged.

Two additional points should be made about part (ii). First, because we allow for the possibility that $\xi'^{\nu} = \xi^{\nu}$ for all $\nu \in \mathcal{N}$, **IND** guarantees the uniqueness of the labour received by agents at any given allocation, arguably an important property. Second, from a formal viewpoint, in the case with $c'^{\nu} \neq c^{\nu}$ for some $\nu \in \mathcal{N}$, the condition is very weak since it hinges on the existence, for a given economy, of multiple equilibria with the same price vector in which agents make the same choices concerning labour and production activity.

We are ready to derive our main characterisation result which identifies a unique class of definitions of labour exploitation $d \in \mathcal{D}_L$ which satisfy **RE**, **SI**, and **IND** in all economies and at all equilibria with positive profits.

To be precise, let $\mathcal{D}_L^{\tau} \subseteq \mathcal{D}_L$ be the set of definitions of labour exploitation such that for all $E \in \mathcal{E}$ and $\left((p, w), (\xi^{\nu})_{\nu \in \mathcal{N}}\right) \in \mathcal{RS}_E$ with $\pi > 0$, $\alpha_l^{\bar{c}^{\nu}} = \alpha_l^{\underline{c}^{\nu}} = \frac{pc^{\nu}}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)} (\alpha_l^{p,w} + \beta_l^{p,w})$ for all $\nu \in \mathcal{N}$.[25] Then:

Theorem 1. *A definition of labour exploitation $d \in \mathcal{D}_L$ satisfies **RE, SI**, and **IND** if and only if $d \in \mathcal{D}_L^{\tau}$.*

Proof. See Appendix C. □

[25]The latter ratio is well defined because $p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right) > 0$ at an RS.

Theorem 1 provides strong support to the so-called 'New Interpretation' (Duménil 1980, 1984; Foley 1982, 1986), as recently extended to individual agents by Yoshihara (2010) and Veneziani and Yoshihara (2015, 2017), which can be formalised as follows.[26]

Definition 4. Let $E \in \mathcal{E}$ and $\left((p, w), (\xi^\nu)_{\nu \in N}\right) \in \mathcal{RS}_E$. For any $\nu \in \mathcal{N}$, let $\tau^{c^\nu} \in [0,1]$ be such that $\tau^{c^\nu}\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right) \in \mathcal{B}(p, c^\nu)$. Agent ν is *exploited* if and only if $\Lambda^\nu > \tau^{c^\nu}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$ and an *exploiter* if and only if $\Lambda^\nu < \tau^{c^\nu}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$.

The normative intuition captured by Definition 4 is quite simple: an agent is exploited (an exploiter) if and only if the share of aggregate labour contributed by the agent in productive activities is greater (lower) than the share of aggregate income received. In other words, it embodies a notion of proportionality between contribution and reward that has been explored in normative economics (Roemer and Silvestre 1993; Moulin 1990) and whose philosophical roots can be traced back to Aristotle (Maniquet 2002) and Kant (Roemer 2019).

According to Theorem 1, Definition 4 is the only definition in the literature that satisfies a small set of theoretically robust and formally weak axioms that capture some widely shared intuitions. Actually, Theorem 1 has a rather striking implication: because *all* of the main definitions of labour exploitation satisfy **SI** and **IND**,[27] it is the rather mild axiom **RE** that rules out all alternative approaches except Definition 4. In other words, the main distinguishing feature of Definition 4 is that it is the only approach that guarantees that whenever some agent is identified as exploited, there is someone exploiting, and vice versa. This is arguably a foundational property in exploitation theory and yet, as it turns out, if Morishima's (1974) celebrated definition is adopted, for example, it is not difficult to construct examples in which **RE** is violated and *all* agents are exploited.

Perhaps more importantly, however, Theorem 1 provides robust theoretical foundations to an approach that – unlike many of the definitions in the literature – identifies the exploitation of individual agents based on empirically measurable magnitudes. In the next sections we take this approach to the data.

[26]Intuitively, for any agent $\nu \in \mathcal{N}$, τ^{c^ν} represents ν's share of national income, and so $\tau^{c^\nu}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$ is the share of social labour that ν receives by earning income barely sufficient to buy c^ν. Then, as in the New Interpretation, the notion of exploitation is related to the production and distribution of national income and social labour.

[27]See Appendix B for a formal proof.

5 Measuring exploitation at the individual level

Theorem 1 identifies the definition of exploitation that satisfies the three axioms **RE**, **SI** and **IND**. This definition can be used to derive an empirical measure of exploitation intensity at the individual level. To see this, let us introduce time explicitly and consider an economy observed at t. Let I_t^ν and I_t denote, respectively, income received by agent ν and aggregate income. In our framework and notation, we have $I_t^\nu = p_t c_t^\nu$ and $I_t = p_t \left(\widehat{\alpha}_t^{p,w} + \widehat{\beta}_t^{p,w} \right)$.

The definition identified by Theorem 1 implies that an agent is exploited if and only if $\frac{\Lambda_t^\nu}{\Lambda_t} > \frac{I_t^\nu}{I_t}$, and an exploiter if and only if $\frac{\Lambda_t^\nu}{\Lambda_t} < \frac{I_t^\nu}{I_t}$. The intuition is straightforward: an agent is exploited if their share in total labour contribution is higher than their share in total income, and an exploiter if the opposite is true.

We define *exploitation intensity*, ε_t^ν, as the logarithmic distance of an individual from the exploitation-neutral point (i.e., the distance from the threshold that divides exploiters and exploited), such that $\varepsilon_t^\nu > 0$ means that ν is exploited, while $\varepsilon_t^\nu < 0$ identifies ν as an exploiter, and $\varepsilon_t^\nu > \varepsilon_t^{\nu'}$ means that agent ν is more exploited than agent ν'. Formally:

$$\varepsilon_t^\nu = \ln\left(\frac{\Lambda_t^\nu}{\Lambda_t}\right) - \ln\left(\frac{I_t^\nu}{I_t}\right). \tag{5}$$

The index ε_t^ν is therefore a continuous variable that measures the *extent* to which an individual is exploited as the logarithmic difference between an individual's share of labour contribution and their share of income. For example, an exploitation intensity index of +0.1 means that the individual's share of labour is approximately 1.1 times their share of income.

In view of empirical estimation, it is convenient to express exploitation intensity equivalently in terms of *labour-income ratios*. We thus rearrange equation 5 as follows

$$\varepsilon_t^\nu = \ln\left(\frac{\Lambda_t^\nu}{I_t^\nu}\right) - \ln\left(\frac{\Lambda_t}{I_t}\right). \tag{6}$$

Seen in this way, the exploitation intensity index ε_t^ν is approximately equal to the percent deviation of the individual labour-income ratio from the economy-wide ratio.

In the remainder of this paper we use nationally representative survey data to estimate the distribution of ε_t^ν in the United States. We focus on pre-tax and transfers measures, thus reflecting market relations and not taking into account direct government redistribution.

5.1 Measuring effective labour performed

It is not immediately obvious how to measure empirically *effective* labour performed, Λ^ν. For while labour hours are observable, individual skills and productive contributions are not.

In a neoclassical perfectly competitive economy, wage differentials would reflect differences in productivity, with all individuals receiving the same wage per unit of effective labour. Under this assumption, the observed hourly wage is $w^\nu = \bar{w}s^\nu$, where $\bar{w}$ is the (constant) wage per unit of effective labour, and effective labour can be measured as $\Lambda^\nu = \frac{w^\nu \lambda^\nu}{\bar{w}}$, where $\bar{w}$ can be ignored since it is a constant common to all individuals.

Nonetheless, actual economies are very far from the perfectly competitive benchmark and assuming that wages closely mirror productive contributions seems unrealistic, as it leaves no space for any kind of rent, social privilege, discrimination or other structural social factors that are key to an empirically-relevant conception of exploitation. A rich recent literature on labor market power has found that wage markdowns, defined as negative deviations of wages from marginal productivity, are substantial and vary widely across firms.[28] Moreover, it is well established that observed gender and racial wage gaps cannot be entirely attributed to productivity differentials.

Alternatively, one could assume that we can observe and measure a vector $\mathbf{x}$ of variables that fully determine productivity and their effect on wages. Assume $w^\nu = s^\nu \psi^\nu$, $E(\psi^\nu | s^\nu) = E(\psi^\nu) = 1$, and $s^\nu = f(\mathbf{x})$, where ψ is a multiplicative error term. One could then use the standard Mincerian approach to estimate effective labour as $\hat{\Lambda}^\nu = \hat{f}(\mathbf{x})\lambda$, where $\hat{f}()$ is an econometric estimate of $f()$ from sample data on w and $\mathbf{x}$.

A measure based on Mincer-equations, however, would be based on very strong assumptions about the determinants of productivity and about them being uncorrelated with other factors affecting wages (the error term ψ needs to be independent of the vector of characteristics affecting productivity). In practice, it is not easy to interpret because it is a mix of differences in income, work time, education and experience levels, returns to education and experience, and potential biases in estimating the latter.

A third possible approach consists in assuming that everyone's labour is in some relevant sense equally valuable, and normalising skills to one, so that effective labour performed is equal to hours worked ($\Lambda^\nu = \lambda^\nu$ in our notation). Clearly, this approach relies on strong assumptions too: individual skills and work effort are unlikely to be identical.

In light of the limitations just discussed, we believe that, *on its own*, none of the three

[28] For example, Yeh, Macaluso, and Hershbein (2022) find that in the US the average worker earns only 65 cents on the marginal dollar generated, and that there is substantial variability in markdowns even within the same narrowly-defined industry.

methods discussed can provide an exact description of exploitative relations in advanced economies. *Taken together*, however, they are likely to provide a reasonably accurate picture of exploitation. For assuming that wages fully reflect productive contributions (as in the first approach) or that they do not correlate with skills at all (as in the third approach) can be seen as estimating, respectively, a lower and an upper bound to exploitative relations.

Therefore we have estimated the distribution of the exploitation intensity index using all three approaches. In the main text, however, for reasons of space, we present the estimates under the assumption that effective labour equals hours worked. We believe that the third approach yields the most relevant and readily interpretable empirical index of exploitation.

First, empirically, this approach is the most robust, as it only requires observing hours worked, an observable variable available in nationally representative surveys. Second, it is based on simple labour-income ratios, which have a straightforward interpretation in terms of hours worked per dollar of market income. Third, the third approach yields some independently interesting normative insights, since it captures both income and hours worked, two crucial determinants of well-being.

Nonetheless, the key conclusions of our empirical analysis are robust: the results obtained adopting the other two approaches are presented in Appendices E and F, respectively, and briefly described in Section 6.5.

5.2 Construction of the exploitation index from survey data

We use the nationally representative Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) to estimate the distribution of the exploitation intensity index, ε_t^ν, in the US population. While the CPS ASEC starts in 1962, sufficient information to compute individual hours worked and total income is not available before 1975. Our sample thus covers the 1975-2022 period.[29]

For each individual ν in the CPS who is over 14 years old, we compute the number of hours of work during year t, Λ_t^ν, and total market income, I_t^ν: Λ_t^ν is the number of weeks worked times the usual hours worked per week, while I_t^ν is calculated as the sum of wages,

[29]The CPS is administered by the US Census Bureau and was retrieved from the International Public Use Microdata Series (Flood, King, Rodgers, Ruggles, Warren, and Westberry 2021). The CPS ASEC asks information about income and hours worked in the previous year, therefore the 1975-2022 period corresponds to the 1976-2023 editions of the survey. This information is only available for individuals who are at least 15 years old (14 in pre-1979 editions)

business/professional practice/farm income, rental income, interest and dividends.[30] [31]

We apply an adjustment for workers who perform significant unpaid work in a business owned by their family: 'unpaid family workers' report unrealistically high labour-income ratios, since they probably benefit from the income that they generate for their family, while not reporting it as their own. At the same time, and for the same reason, the labour-income ratios of their family members tend to be underestimated. To adjust for that, we attribute to unpaid family workers and to all their family members the overall labour-income ratio of their household.

It is well documented that nonwage income sources are severely under-reported in survey data, including in the CPS (Rothbaum 2015). Following Hoffmann et al. (2020), we scale up capital incomes using the under-reporting ratios computed by Rothbaum (2015) based on the discrepancy between CPS data and national income and product accounts.[32]

We remove from the sample observations for which either the overall labour-income ratio or the ratio of work and personal business income over hours worked imply hourly earnings below two-thirds of the prevailing federal minimum wage, employees whose wage/hours ratio implies a hourly salary above 50,000 1999 dollars (≈ 81,300 2022 dollars), and those with zero hours worked but positive wage income. These observations are very likely to reflect reporting mistakes.[33]

Individuals who report no market income are excluded from the analysis, as their labour-income ratio and exploitation status is undetermined, both mathematically and conceptually.

[30] This is a 'minimal' definition of market income. Due to data limitations, it does not include realised capital gains, private pension payments, alimony, scholarships from private entities, imputed housing services, and other items that are included in the US Census definition (https://www.census.gov/programs-surveys/cps/data/data-tools/cps- table-creator-help/income-definitions.html).

[31] In order to avoid the identification of individuals with extremely high incomes, CPS income data are topcoded: incomes above a chosen threshold are not reported in the same way as they were declared by respondents. The topcoding procedures employed in the CPS have improved over time (`https://cps.ipums.org/cps/topcodes_tables.shtml`). Since 2011, a rank proximity swapping procedure is used, which ensures that the distribution of income is preserved even above the topcoding threshold, while guaranteeing anonymity. To keep income data consistent over time, and preserve the distribution of top incomes in all years in our sample, we use the retrospective revised income topcoding files provided by the US Census Bureau and made available by Flood et al. (2021). Thus, we apply to all observations in all years in our sample the same rank proximity swapping procedure used by the US Census Bureau since 2011.

[32] Specifically, we multiply personal business income by (1/0.356); interest income by 1/0.675; dividend income by 1/0.695; rental income by 1/0.274. While this adjustment is important, two limitations should be acknowledged. First, available information only allows to estimate and apply *fixed* under-reporting ratios, thus neglecting variation over time in the extent of under-reporting. Second, as noticed by Hoffmann et al. (2020), it seems likely that the extent of under-reporting is larger in the upper part of the distribution, but only *aggregate ratios* are currently available. (Rothbaum (2015) does not find any significant under-reporting of wage and salary earnings, and thus we do not adjust those figures.)

[33] In some cases, these abnormally high labour-income ratios can also arise from negative personal business or rental income.

Individuals who report zero hours worked and non-negligible market (nonwage) income, instead, pose a challenge: theoretically, they should be classified as exploiters, but their exploitation index is undetermined given the logarithmic specification in Equation 6. We deal with these observations as follows. Individuals who do not work because they are retired, homemakers, or students are removed from the sample, as they do not yet (or anymore) participate meaningfully in market activities. As for the remaining observations with zero hours worked and non-negligible market income we attribute them 1 hour of work in the year, so that their exploitation index can be computed.[34]

The resulting sample includes 4,025,201 observations during 1975-2022 with a well-defined exploitation intensity index. The distribution of ε_t^{ν} for selected years is shown in Figure 1.

6 Exploitation in the United States, 1975-2022

Using our main dataset, based on ASEC CPS data, we examine exploitation intensity in the US in 1975-2022. Sections 6.1 and 6.2 detail the dynamics of labour-income ratios and ε_t^{ν} by percentile over the sample period. Section 6.3 analyses inequality in exploitation, and compares it with income inequality. In Section 6.4, we assess gender, racial and socioeconomic predictors of exploitation. Section 6.5 discusses results using alternative measures of labour contribution (presented in more detail in Appendices E and F).

6.1 Labour-income ratios

Given that ε_t^{ν} is based on individual labor-income ratios, we start by briefly reviewing their dynamics by percentile in Figure 2. The left panel displays levels; its unit of measure is hours worked per (constant 1999) dollar of market income. In the right panel, 1975 values are normalised to 100 in order to display changes over the sample period. Recall that a lower labour-income ratio implies less exploitation. Therefore the bottom percentiles are either exploiters or less intensely exploited, while the top percentiles are the most exploited.

Over the 1975-2022 period labour-income ratios went down substantially for the lower percentiles, while the higher percentiles (the most exploited) display only limited improvements: the average number of hours of work per dollar of total market income has decreased by 28 percent in real terms for the 10th percentile, but only by 7 percent for the 90th per-

[34] We define 'nonnegligible' market income as at least 7,000 1999 dollars ($\approx$ 11,400 2022 dollars) per year. This is a very limited number of observations (around 2,000 in the whole sample period, or 0.5% of the overall sample) and removing them from the analysis does not meaningfully alter any results.

centile and 10 percent for the median American. The divergence seems to have originated in the 1980-1989 period, when labour-income ratios decreased substantially for the lowest percentiles, while *increasing* for the median American and for the higher percentiles.

6.2 Exploitation intensity

Figure 3 displays exploitation intensity by percentile over the 1945-2022 period. Figure 4 and Figure 5 zoom in, respectively, on the most exploited half of the population and those with negative exploitation and highlight dynamics over the period.

Figure 3 (along with the histograms of Figure 1) suggests that most people are exploited: the fraction of the population with $\varepsilon_t^\nu > 0$ hovers slightly above 75%, with no clear trend, over the period. For the median American, ε_t^ν is around 0.50 in 2022. Given the interpretation of logarithmic differences as approximate percentage changes, this implies a labour-income ratio approximately 50 percentage points above the economy-wide one.

The 1975-2022 period has seen a marked increase in the gap between the top and bottom percentiles of ε_t^ν . Exploitation intensity went up for the most exploited and for the median American, and became even more negative for those identified as exploiters. This is seen most clearly in Figures 4 and 5: ε_t^ν increased by almost 30 percent for the median American, almost 15 percent for the 90th percentile and over 33 percent for the 99th percentile; but it decreased by over 20 percent per the least exploited 10% of the population.

6.3 Inequality in exploitation

To quantify the increase in inequality in exploitation and compare it with the rise in income inequality, this subsection computes inequality measures for the exploitation intensity index.

Figure 6 shows that the standard deviation of labour-income ratios increased substantially in the last four decades, due to a large increase between 1980 and 1989.

Figure 7 displays the Gini index for labour-income ratios (red line) and, for the sake of comparison, for market income (grey line). The Gini index for labour-income ratios increased markedly – by around 0.10 points, which amounts to a 22 percent increase – over the sample period. By way of comparison, the Gini index for income inequality increased by 1.6 percent in the whole sample and 19 percent among full-time workers.[35]

[35] We include this figure because in the literature income inequality is usually measured by restricting the analysis to full-time workers. Note that it would not make sense to compute inequality in labour-income ratios only among full-time workers: by definition the latter work the same number of hours, therefore variation in labour-income ratios can only come from changes in income.

6.4 Gender, racial and socioeconomic predictors of exploitation

Which groups are most exploited according to our index? To answer this question, in Table 1 we regress exploitation intensity on a set of binary variables capturing potentially relevant individual characteristics, controlling for a full set of age-by-year fixed effects.

We find that women and blacks tend to be substantially more exploited. The gender exploitation gap is around 22 percent in the whole sample. While currently still substantial, it has decreased significantly over the sample period: from 28 percent in 1975-1989 to 18 percent in 2005-2022. The racial exploitation gap is around 10 percent over the whole period. It has decreased from 12 percent in 1975-1989 to 8 percent in 1990-2004, but then increased again, to 10 percent, in 2005-2022.

Among socio-economic groups, rentiers (defined as individuals who receive positive capital income), entrepreneurs and the retired tend to be substantially less exploited than individuals whose income only consists of wages. The unemployed, instead, tend to be more exploited, by around 7 percent in the whole sample period.[36]

6.5 Alternative measures of effective labour performed

While our main analysis uses hours worked as the measure of labour contribution, Appendices E and F display results using alternative approaches. In Appendix E we use earned income as the measure of labour contribution, under the assumption that wage differentials reflect differences in labour productivity. This 'earned income' approach produces a 'lower bound' measure for exploitation intensity. In Appendix F, we use Mincer-type regressions in order to separate the component of earned income that reflects skills (and therefore productivity) from the component that reflects other factors unrelated to productivity ('Mincer-equation' approach).[37]

By construction, the resulting indices suggest lower levels of exploitation intensity – substantially lower in the case of the 'earned income' approach but only modestly lower in the case of the 'Mincer-equation' approach (Figures E.2 and F.3). However, the dynamics are broadly similar: exploitation intensity increases for the most exploited half of the population, and decreases for the better-off, and inequality in exploitation increased more than income inequality (Figures E.3, F.6 and F.7).

[36]In interpreting these results, it should be noted that only retired and unemployed who work are included in our analysis (see Section 5). Moreover, public pensions and unemployment benefits do not enter our measure of exploitation, since we focus on *market* income.

[37]See Section 5.1 and Appendices E and F for a more detailed discussion of both approaches.

7 The relevance of exploitation intensity

A natural question concerns the *relevance* of the index we are proposing: can exploitation contribute to explain relevant socio-economic outcomes? Does the exploitation index possess *independent* explanatory power? Due to space and, partly, data limitations, we cannot provide a comprehensive answer to this question here. Nonetheless, in closing this paper, we briefly provide some preliminary results suggesting that exploitation intensity does matter, in the sense of carrying predictive power for important socio-economic outcomes at the individual level *even after controlling for income.*

Unfortunately, the CPS contains limited information concerning various socio-economic outcomes of interest and therefore in this section we use the data derived in the General Social Survey (GSS). GSS data provide rougher estimates of individual labour and income and smaller samples, but include information on well-being, class identification and political opinions. Appendix G describes the GSS and explains how we build ε_t^{ν} from this dataset,[38] as well as binary variables for whether an individual reports to be happy, healthy, satisfied with their job, a member of the working class, and a Democratic voter.

Figure 8 displays average individual outcomes by percentiles of exploitation intensity. More exploited individuals report lower happiness, worse health, lower job satisfaction and are more likely to identify as working class. They are more likely to support the Democratic party in the 1974-1995 period but not thereafter.

Tables 2 and 3 report results from OLS regressions of each socio-economic outcome on percentile of exploitation intensity and a full set of age and year fixed effects, with and without controlling for percentile of income. The specifications that control for income (columns 2, 4 and 6 of both Tables) compare individuals within the same percentile of income (as well as of the same age and interviewed in the same year), but in a different percentile of exploitation intensity. The regression results confirm the insights from Figure 8 and additionally suggest that the relation between exploitation and these outcomes is robust to controlling for income (with the sole exception of job satisfaction).

The size of these correlations is non-negligible, and roughly comparable to the size of the correlation between these variables and income. For example, moving up by 10 percentiles in the distribution of exploitation reduces by 1.3 percent the probability of feeling very happy (0.5 percent when controlling for income percentile) and increases by 6 percent the probability of identifying as working class (3.5 percent when controlling for income percentile).

[38] In this section, we focus only on the simple measure of exploitation intensity and do not consider either the Mincer equation approach or the earned income approach.

Figure 9 shows the evolution of the association between exploitation and political affiliation. It displays the effect of a 1 percentile increase in exploitation intensity on the probability of being a Democrat and how it changed over the entire sample period. For comparison, in the right panel the same exercise is done for income. Both income and exploitation lose their predictive power since the mid-1990s, consistent with previous analyses (for example, Gethin, Martínez-Toledano, and Piketty (2022)).

8 Conclusions

In this paper we have provided a comprehensive analysis – both theoretical and empirical – of the concept of exploitation as the unequal exchange of labour. Theoretically, we have derived the first axiomatic characterisation of a definition of exploitation in a general equilibrium framework. Contrary to the received view, a unique definition of exploitation can be identified that is logically consistent and incorporates a small set of normatively robust properties. Empirically, we have shown that this definition can be used to formalise an index that can be used to measure exploitation intensity at the individual level. Contrary to a widespread belief, far from being metaphysical, the concept of exploitation provides the foundation for a rich empirical analysis of contemporary economies.

Our empirical results indicate that inequality in exploitation has increased substantially in the US since the early 1980s, and by more than income inequality. The largest increase in exploitation intensity has been experienced by two groups: the most exploited percentile, and those in the middle. We have found a significant gender exploitation gap, which has declined steadily since 1975, but is still substantial (20% on average) in 2005-2022. The racial exploitation gap does not have a clear trend and is around 10% in 2005-2022. A descriptive analysis of exploitation and individual outcomes finds that higher exploitation is associated with lower subjective well-being and health, and a higher probability of voting for the Democratic party before the 1990s, and that these correlations survive even after controlling for income.

By proving the existence of a logically consistent definition of exploitation, uniquely characterized by a small number of normatively relevant properties and subject to empirical estimation, we hope to have moved a first step towards a constructive reappraisal of the concept of exploitation in economics. We believe that the exploitation intensity index proposed here represents a potentially fruitful addition to the economist's toolbox.

References

ACLU (2023). Captive Labor: Exploitation of Incarcerated Workers. Technical report, American Civil Liberties Union (ACLU) and the University of Chicago Law School Global Human Rights Clinic (GHRC).

Alvaredo, F., L. Chancel, T. Piketty, E. Saez, and G. Zucman (2017). Global inequality dynamics: New findings from wid. world. *American Economic Review 107*(5), 404–409.

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

Bartley, T. and C. Child (2014). Shaming the corporation: The social production of targets and the anti-sweatshop movement. *American Sociological Review 79*(4), 653–679.

Bryan, A. (2023). Structural domination and freedom in the labor market: From voluntariness to independence. *American Political Science Review 117*(2), 692–704.

Chan, S. Y. (2024). Expropriation as reparation. *American Journal of Political Science forthcoming*, 1–15.

Desmond, M. and N. Wilmers (2019). Do the poor pay more for housing? exploitation, profit, and risk in rental markets. *American Journal of Sociology 124*(4), 1090–1124.

Duménil, G. (1980). *De la valeur aux prix de production*. Paris: Economica.

Duménil, G. (1984). The So-Called 'Transformation Problem' Revisited: A Brief Comment. *Journal of Economic Theory 33*(2), pp. 340–348.

Ferguson, B. (2021). Are we all exploiters? *Philosophy and Phenomenological Research 103*, pp. 535–546.

Flaschel, P. (1983). Actual Labor Values in a General Model of Production. *Econometrica 51*(2), pp. 435–454.

Flaschel, P. (2010). *Topics in Classical Mirco- and Macroeconomics*. New York, NY: Springer.

Fleurbaey, M. (1995). Three solutions for the compensation problem. *Journal of Economic Theory 65*(2), 505–521.

Fleurbaey, M. (1996). *Théories Économiques de la Justice*. Paris: Economica.

Fleurbaey, M. (2014). The facets of exploitation. *Journal of Theoretical Politics 26*(4), pp. 653–676.

Fleurbaey, M. and F. Maniquet (1999). Cooperative production with unequal skills: The solidarity approach to compensation. *Social Choice and Welfare 16*(4), 569–583.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry (2021). Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]. https://doi.org/10.18128/D030.V9.0.

Foley, D. K. (1982). The Value of Money the Value of Labor Power and the Marxian Transformation Problem. *Review of Radical Political Economics 14*(2), pp. 37–47.

Foley, D. K. (1983). Realization and Accumulation in a Marxian Model of the Circuit of Capital. *Journal of Economic Theory 28*(2), pp. 300–319.

Foley, D. K. (1986). *Understanding Capital*. Cambridge, MA: Harvard University Press.

Galanis, G., R. Veneziani, and N. Yoshihara (2019). The dynamics of inequalities and unequal exchange of labor in intertemporal linear economies. *Journal of Economic Dynamics & Control 100*, pp. 29–46.

Gerver, M. (2022). The case for permanent residency for frontline workers. *American Political Science Review 116*(1), 87–100.

Gethin, A., C. Martínez-Toledano, and T. Piketty (2022). Brahmin Left Versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948–2020. *The Quarterly Journal of Economics 137*(1), pp. 1–48.

Gourevitch, A. (2018). The right to strike: A radical view. *American Political Science Review 112*(4), 905–917.

Hawkins, J. S. and E. J. Emanuel (2008). *Exploitation and Developing Countries: The Ethics of Clinical Research*. Princeton: Princeton University Press.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Chapter 7 earnings functions, rates of return and treatment effects: The mincer equation and beyond. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, pp. 307–458. Amsterdam: Elsevier.

Hoffmann, F., D. S. Lee, and T. Lemieux (2020, November). Growing income inequality in the united states and other advanced economies. *Journal of Economic Perspectives 34*(4), 52–78.

Horton, J. (2019). The exploitation problem. *Journal of Political Philosophy 27*, pp. 469–479.

Howard, S. (2023). The surrogacy trade: proliferating bans and an opportunistic industry raise a worrying health risk. *British Medical Journal 383*, pp. 1–4.

ILO (2024). Fair recruitment Roadmap. A guide for national action. Technical report, International Labour Organization.

Kuznets, S. (1954). Shares of upper income groups in income and savings. Technical report, National Bureau of Economic Research.

Labour Party (2024). Change. Labour Party Manifesto 2024. London.

Lamkin, M. and C. Elliott (2018). Avoiding Exploitation in Phase I Clinical Trials: More than (Un)Just Compensation. *Journal of Law, Medicine & Ethics 46*, pp. 52–63.

Lee, J. Y. (2023). Surrogacy: beyond the commercial/altruistic distinction. *Journal of Medical Ethics 49*(3), pp. 196–199.

Maniquet, F. (2002). A study of proportionality and robustness in economies with a commonly owned technology. *Review of Economic Design 7*, pp. 1–15.

Maniquet, F. and Y. Sprumont (2004). Fair production and allocation of an excludable nonrival good. *Econometrica 72*(2), pp. 627—-640.

McGee, D. (2025). Exploitation Through Racialization. *Quarterly Journal of Economics*, forthcoming.

Mears, A. (2015). Working for free in the vip: Relational work and the production of consent. *American Sociological Review 80*(6), 1099–1122.

Milanovic, B. (2016). *Global inequality: A new approach for the age of globalization.* Harvard University Press.

Morishima, M. (1974). Marx in the Light of Modern Economic Theory. *Econometrica 42*(4), pp. 611–632.

Moulin, H. (1987). Egalitarian-equivalent cost sharing of a public good. *Econometrica 55*(4), 963–976.

Moulin, H. (1990). Joint Ownership of a Convex Technology: Comparison of Three Solutions. *Review of Economic Studies 57*, pp. 439–452.

Pazner, E. A. and D. Schmeidler (1978). Egalitarian equivalent allocations: A new concept of economic equity. *Quarterly Journal of Economics 92*(4), pp. 671–687.

Piketty, T. and E. Saez (2003, 02). Income inequality in the united states, 1913–1998*. *The Quarterly Journal of Economics 118*(1), 1–41.

Piketty, T., E. Saez, and G. Zucman (2018). Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics 133*(2), 553–609.

PSOE (2021). Ponencia Marco. 40 Congreso PSOE. Madrid.

Roemer, J. E. (1980). A General Equilibrium Approach to Marxian Economics. *Econometrica 48*(2), pp. 505–530.

Roemer, J. E. (1981). *Analytical foundations of Marxian economic theory.* Cambridge: Cambridge University Press.

Roemer, J. E. (1982). *A General Theory of Exploitation and Class.* Cambridge, MA: Harvard University Press.

Roemer, J. E. (2019). *How we cooperate.* New Haven: Yale University Press.

Roemer, J. E. and J. Silvestre (1993). The proportional solution for economies with both private and public ownership. *Journal of Economic Theory 59*, pp. 426–444.

Rothbaum, J. L. (2015). Comparing Income Aggregates: How do the CPS and ACS Match the National Income and Product Accounts, 2007-2012. SEHSD Working Paper 2015-01.

Sauer, C., P. Valet, S. Shams, and D. Tomaskovic-Devey (2021). Categorical distinctions and claims-making: Opportunity, agency, and returns from wage negotiations. *American Sociological Review 86*(5), pp. 934–959.

Smith, T. W., M. Davern, J. Freese, and S. L. Morgan (2021). General social surveys, 1972-2018 [machine-readable data file]. Chicago: NORC, 2019.

SPD (2021). The SPD's programme for the future. Berlin.

Steiner, H. (1984). A liberal theory of exploitation. *Ethics 92*, pp. 225–241.

Stiglitz, J. E., A. Sen, and J.-P. Fitoussi (2009). Report by the commission on the measurement of economic performance and social progress.

Tan, S. H. (2020). Surrogacy and human flourishing. *Journal of Legal Philosophy 45*, pp. 49–79.

Temin, D. M. (2024). Wages for earthwork. *American Political Science Review*, 1–14.

Thomson, W. (1994). Notions of equal, or equivalent, opportunities. *Social Choice and Welfare 11*(2), pp. 137–156.

Tomaskovic-Devey, D., M. Hällsten, and D. Avent-Holt (2015). Where do immigrants fare worse? modeling workplace wage gap variation with longitudinal employer-employee data. *American Journal of Sociology 120*(4), 1095–1143.

Valdez, I. (2020). Reconceiving immigration politics: Walter benjamin, violence, and labor. *American Political Science Review 114*, 95–108.

Veneziani, R. (2007). Exploitation and time. *Journal of Economic Theory 132*(1), pp. 189–207.

Veneziani, R. and N. Yoshihara (2012). Profits and Exploitation: A Reappraisal. *Advances in Mathematical Economics 16*, pp. 85–109.

Veneziani, R. and N. Yoshihara (2015). Exploitation in economies with heterogeneous preferences, skills and assets: An axiomatic approach. *Journal of Theoretical Politics 27*(1), pp. 8–33.

Veneziani, R. and N. Yoshihara (2017). Globalisation and Inequality in a Dynamic Economy: An Axiomatic Analysis of Unequal Exchange. *Social Choice and Welfare 49*, pp. 445–468.

Veneziani, R. and N. Yoshihara (2018). The theory of exploitation as the unequal exchange of labour. *Economics and Philosophy 34*, pp. 381–409.

Vrousalis, N. (2013). Exploitation, Vulnerability, and Social Domination. *Philosophy and Public Affairs 41*, pp. 131–157.

Vrousalis, N. (2022). *Exploitation as domination*. Oxford: Oxford University Press.

Wodtke, G. T. (2016). Social class and income inequality in the united states: Ownership, authority, and personal income distribution from 1980 to 2010. *American Journal of Sociology 121*(5), 1375–1415.

Yeh, C., C. Macaluso, and B. Hershbein (2022, July). Monopsony in the us labor market. *American Economic Review 112*(7), 2099–2138.

Yoshihara, N. (2010). Class and exploitation in general convex cone economies. *Journal of Economic Behavior & Organization 75*(2), pp. 281–296.

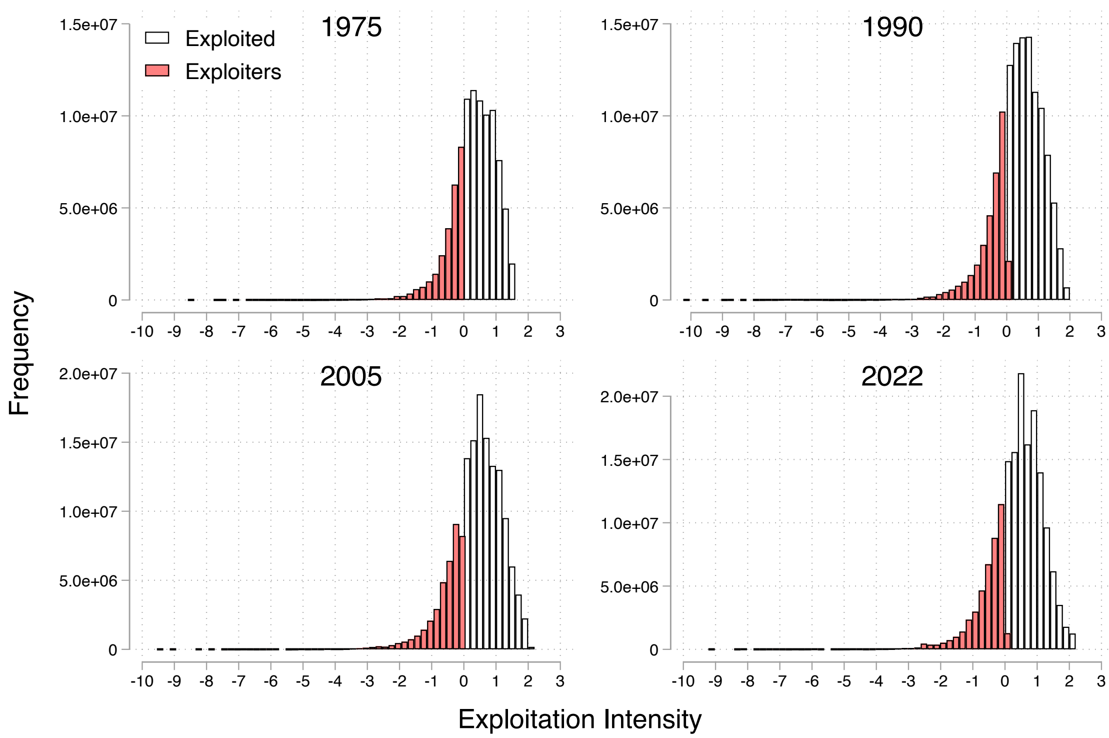
Yoshihara, N. and R. Veneziani (2009). Exploitation as the Unequal Exchange of Labour: An Axiomatic Approach. WP N. 655, Queen Mary, University of London.

Yoshihara, N. and R. Veneziani (2023). The measurement of labour content: an axiomatic approach. *Journal of Economic Behavior & Organization 212*, pp. 392–402.

Zwolinski, M., B. Ferguson, and A. Wertheimer (2022). Exploitation. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2022 ed.). Metaphysics Research Lab, Stanford University.

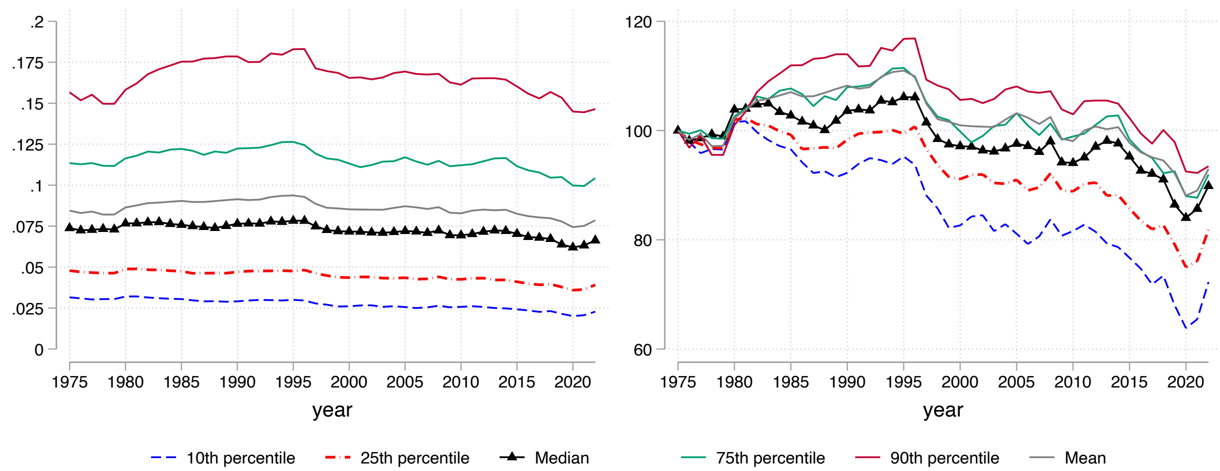
Figures

Figure 1: Distribution of Exploitation Intensity in selected years



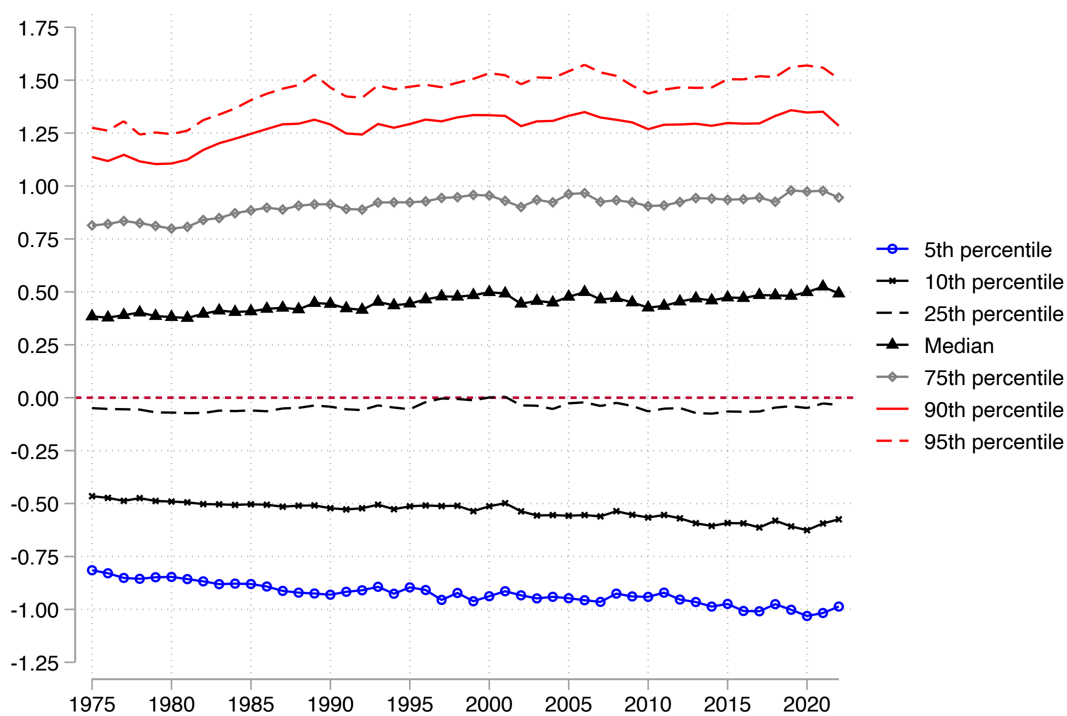
Notes: Estimated from CPS data. Exploitation intensity equals the logarithmic deviation of the individual labour-income ratio from the economy-wide ratio (see Equation (6)*). 'Exploiters' are observations with a individual labour-income ratio below the aggregate ratio. See main text for details.*

Figure 2: Labour-income ratios by percentile (constant 1999 dollars)



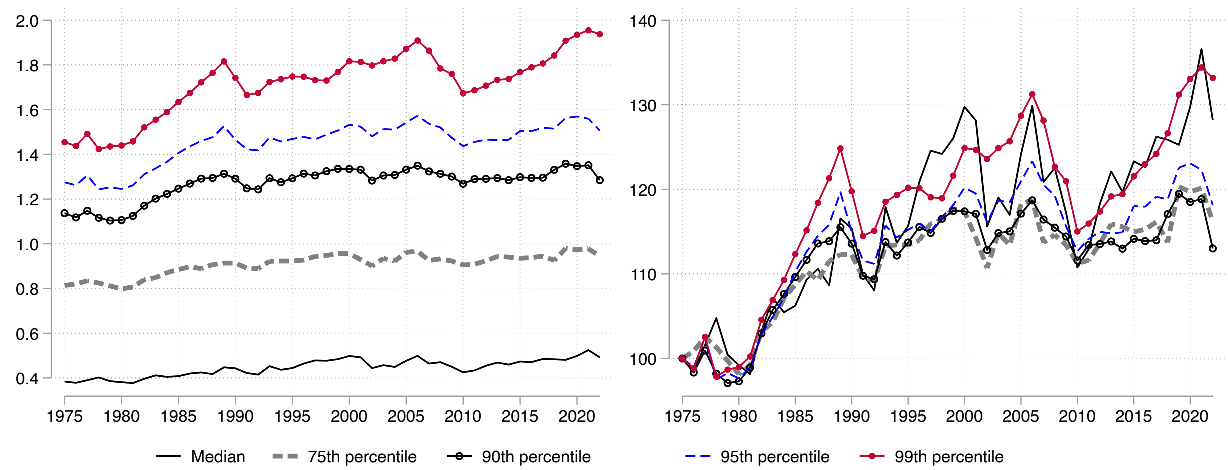
Notes: Estimated from CPS data. labour-income ratios are defined as hours worked during the year divided by total market income (in 1999 constant USD). The left panel displays levels (hours worked per dollar). In the right panel, 1975 values are normalized to 100 to highlight dynamics. See main text for sample and detailed definitions.

Figure 3: Exploitation intensity index by percentile



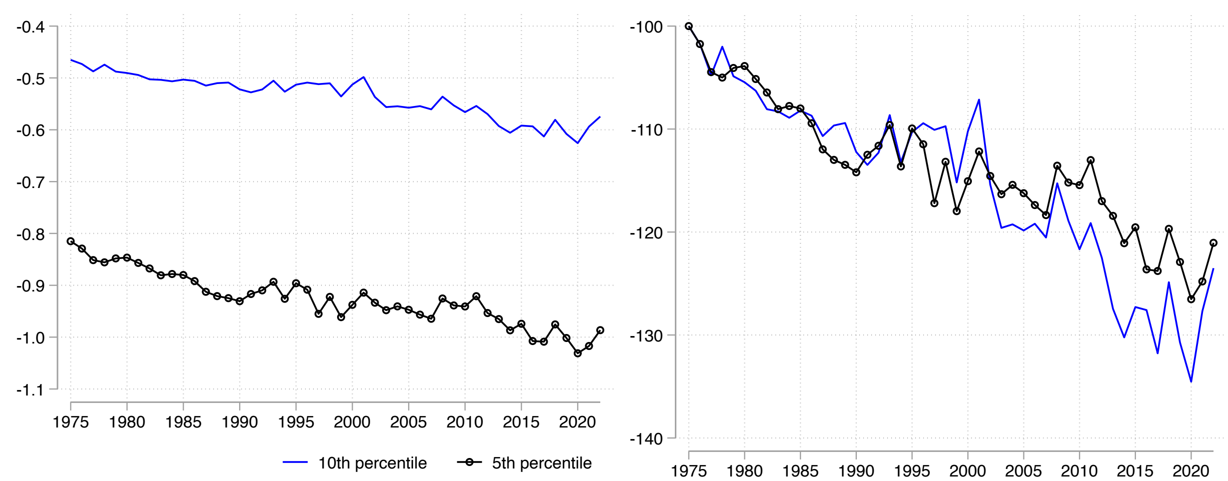
Notes: Estimated from CPS data. Exploitation intensity equals the logarithmic deviation of the individual labour-income ratio from the economy-wide ratio (see Equation (6)*). See main text for details.*

Figure 4: Exploitation intensity for the 50% most exploited



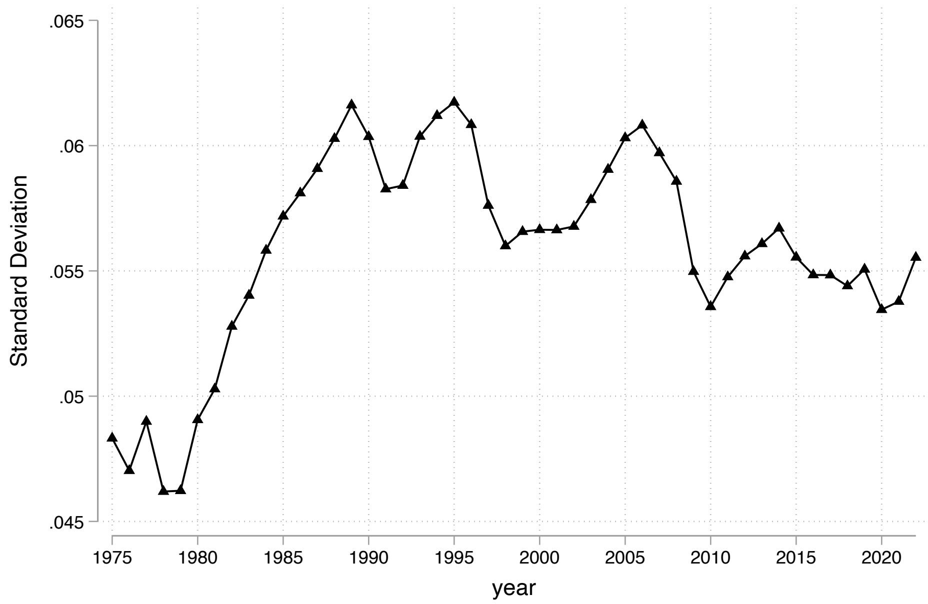
Notes: Estimated from CPS data. Exploitation intensity equals the logarithmic deviation of the individual labour-income ratio from the economy-wide ratio (see Equation (6)*). The left panels displays rates of exploitation. In the right panel, 1975 rates are normalized to 100 to highlight dynamics. See main text for details.*

Figure 5: Exploitation intensity for the 10% least exploited (aka exploiters)



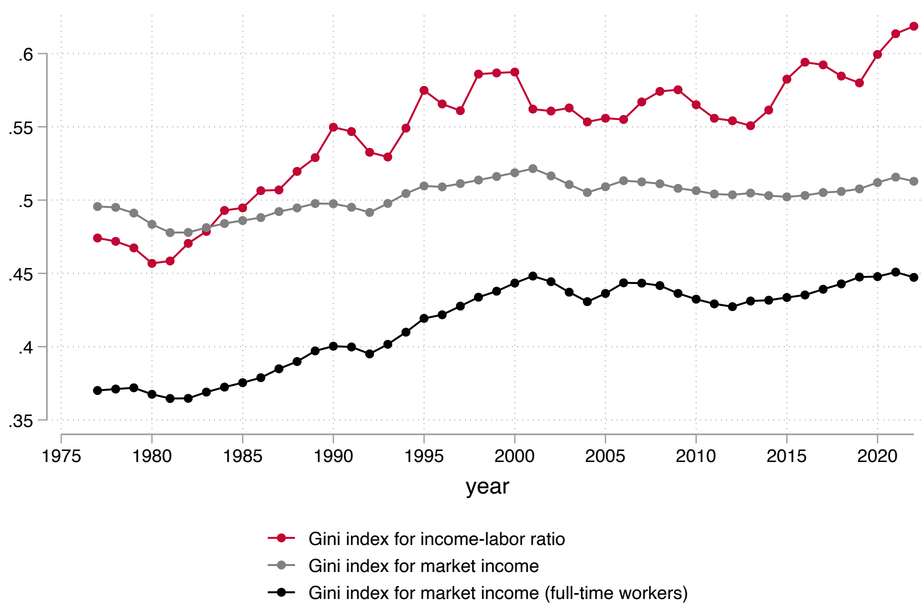
Notes: Estimated from CPS data. Exploitation intensity equals the logarithmic deviation of the individual labour-income ratio from the economy-wide ratio (see Equation (6)*). The left panels displays rates of exploitation. In the right panel, 1975 rates are normalized to -100 to highlight dynamics. See main text for details.*

Figure 6: Inequality: Standard deviation of labour-income ratios (1999 constant US dollars)



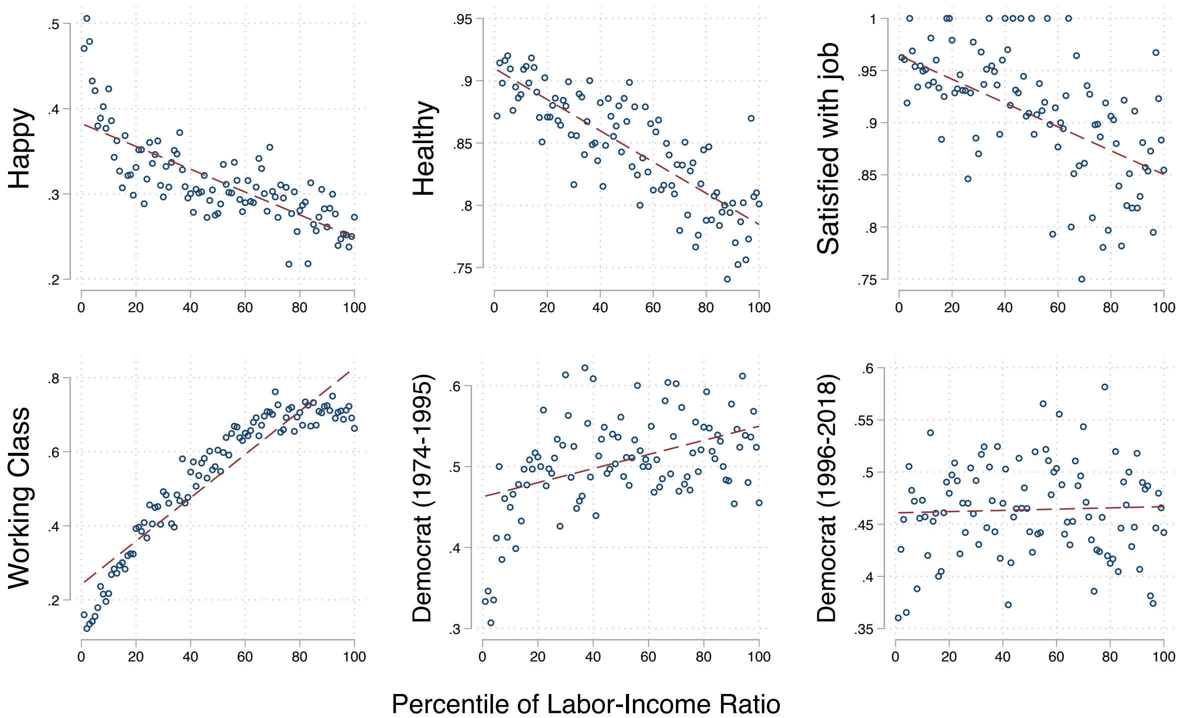
Notes: Estimated from CPS data. Standard deviation of labour-income ratio, measured in constant 1999 dollars. See main text for details.

Figure 7: Inequality: Gini index



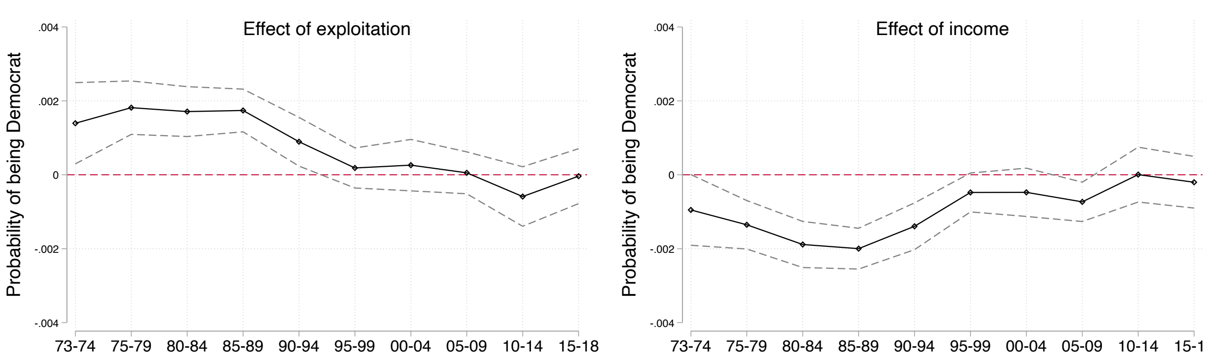
Notes: Estimated from CPS data. Series smoothed with a 3-years backward moving average. Full-time workers are individuals who worked full-time ($\geq$ 35 hours per week) for the full year. See main text for details.

Figure 8: Average individual outcomes by percentile of exploitation intensity



Notes: Estimated from GSS data. Percentile of labour-income ratio (or, equivalently, of exploitation intensity) on the horizontal axis. All outcomes on the vertical axis are binary variables, built as described in Appendix G. Red dashed lines are linear best-fit. See main text and Appendix G for detailed definitions.

Figure 9: Evolution of the association of exploitation and income with political affiliation



Notes: Estimated from GSS data. Effect of a one-percentile increase in exploitation intensity (left panel) or income (right panel) on the probability of supporting the Democratic Party, controlling for age and year fixed effects. See main text and Appendix G for more detail.

Tables

Table 1: OLS estimates for the relation between exploitation and personal characteristics

	(1) Full Sample	(2) 1975-1989	(3) 1990-2004	(4) 2005-2022
female	0.22***	0.28***	0.22***	0.18***
	(0.00)	(0.00)	(0.00)	(0.00)
rentier	-0.44***	-0.37***	-0.45***	-0.47***
	(0.00)	(0.00)	(0.00)	(0.00)
entrepreneur	-0.63***	-0.54***	-0.63***	-0.68***
	(0.00)	(0.00)	(0.00)	(0.00)
retired	-0.14***	-0.14***	-0.11***	-0.15***
	(0.00)	(0.01)	(0.01)	(0.01)
unemployed	0.07***	0.11***	0.07***	0.05***
	(0.00)	(0.01)	(0.00)	(0.00)
black	0.10***	0.12***	0.08***	0.10***
	(0.00)	(0.00)	(0.00)	(0.00)
N	4025042	1131337	1245995	1647710

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is the exploitation intensity index. Estimated from CPS data. All specifications control for a full set of age-by-year fixed effects.

Table 2: Association of exploitation with well-being, health, and job satisfaction.

	(1) Happy	(2) Happy	(3) Healthy	(4) Healthy	(5) Job Satisfied	(6) Job Satisfied
Exploitation	-0.0013***	-0.0005**	-0.0016***	-0.0006**	-0.0009***	-0.0001
	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0003)
Income		0.0010***		0.0013***		0.0010***
		(0.0002)		(0.0002)		(0.0003)
Observations	27416	27416	21529	21529	5533	5533
Adjusted R^2	0.010	0.010	0.027	0.029	0.022	0.024

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimated from GSS data. Exploitation is the percentile in the distribution of exploitation intensity. Income is the percentile in the distribution of income. All outcomes are binary variables, built as described in Appendix G. All specifications include year and age fixed effects.

Table 3: Association of exploitation with class identification and voting behavior

	(1) Working Class	(2) Working Class	(3) Democrat	(4) Democrat	(5) Democrat	(6) Democrat
Exploitation	0.0060***	0.0035***	0.0014***	0.0004	0.0000	-0.0012***
	(0.0001)	(0.0002)	(0.0001)	(0.0003)	(0.0002)	(0.0003)
Income		-0.0031***		-0.0013***		-0.0016***
		(0.0002)		(0.0003)		(0.0003)
Observations	28496	28496	15472	15472	13641	13641
Adjusted R^2	0.119	0.125	0.019	0.020	0.003	0.005

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimated from GSS data. Exploitation is the percentile in the distribution of exploitation intensity. Income is the percentile in the distribution of income. All outcomes are binary variables, build as described in Appendix G. All specifications include year and age fixed effects. Columns (3) and (4) are for the 1973-1995 periods; columns (5) and (6) are for the 1996-2017 period.

Appendix

A Assumptions on the production set

Let $\mathbb{R}_-$ be the set of nonpositive real numbers. The following assumptions on $\mathcal{P} \subset \mathbb{R}^{2n+1}$ hold throughout the paper.

Assumption 1 (A1). For all $\alpha \in \mathcal{P}$, if $\overline{\alpha} \geq \mathbf{0}$ then $\underline{\alpha} \geq \mathbf{0}$ and $\alpha_l > 0$.

Assumption 2 (A2). For all $c \in \mathbb{R}^n_+$, there is a $\alpha \in \mathcal{P}$ such that $\widehat{\alpha} \geqq c$.

Assumption 3 (A3). For all $\alpha \in \mathcal{P}$ and all $\alpha' \in \mathbb{R}_- \times \mathbb{R}^n_- \times \mathbb{R}^n_+$, if $\alpha' \leqq \alpha$ then $\alpha' \in \mathcal{P}$.

A1 implies that both labour and *some* produced input are indispensable to produce any non-negative output vector. **A2** states that any non-negative commodity vector is producible as a net output. **A3** is a standard *free disposal* condition.

B Definitions of exploitation

In this section, some of the main definitions of labour exploitation – suitably extended to economies with heterogeneous skills and general utility functions – are briefly analysed. The purpose is to illustrate the relevance of our axiomatic framework, rather than to provide a comprehensive survey of alternative approaches.

As a starting point, consider a simple economy with a standard Leontief technology (A, L), where A is a square $n \times n$ nonnegative and productive matrix and $L > \mathbf{0}$ is a $1 \times n$ vector describing, respectively, the amount of each input and labour necessary to produce one unit of the n goods. Assume that all agents have equal skills and consume the same subsistence bundle b. Under these assumptions, the definition of labour exploitation is relatively uncontroversial in the literature: the reference bundle is b and the reference amount of labour is equal to vb, where $v = L(I - A)^{-1}$ is the vector of Leontief employment multipliers. Then agent $\nu \in \mathcal{N}^{ted}$ (resp., $\nu \in \mathcal{N}^{ter}$) if and only if the labour she contributes to the economy, Λ^ν, is greater (resp., lower) than the labour she receives, vb.

As soon as these assumptions are dropped, however, the definition of exploitation is not obvious. If more general technologies are considered, the simple generalisation of the standard approach can yield paradoxical results – such as ERBs containing negative amounts of labour – and so various definitions of the labour contained in a *given* bundle have been proposed, focusing either on actual production activities in the economy or on some feasible, possibly counterfactual, technology. Moreover, if agents do *not* consume a given, equal subsistence bundle, then the choice of reference bundle is not obvious: one may focus either on agents' optimal bundles or on some alternative (affordable) bundle.

In his classic definition, Morishima (1974) adopts a counterfactual definition of labour content: for any $c \in \mathbb{R}^n_+$, the labour contained in c is the *minimum* amount of (effective) labour necessary to produce c^ν as net output. Formally,

$$l.v.(c) \equiv \min \{\alpha_l \mid \alpha = (-\alpha_l, -\underline{\alpha}, \overline{\alpha}) \in \phi(c)\}.$$

Further, Morishima (1974) focuses on the agents' optimal bundle, c^{ν}. In the relatively simple models he considers, c^{ν} is uniquely determined, and often equal to the subsistence bundle. In the more general economies considered here, however, an agent's demand correspondence may not be a singleton. In order to take into account this potential indeterminacy, while keeping Morishima's (1974) emphasis on agents' optimal bundles, let $\mathcal{D}^{\nu}(p,w)$ denote agent ν's demand correspondence emerging from MP^{ν} and let $\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})$ represent the projection of $\mathcal{D}^{\nu}$ onto the first n components: at prices (p,w), $\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})$ provides the set of optimal consumption vectors given that the agent supplies Λ^{ν}. Then, Morishima's (1974) definition of exploitation can be extended as follows:[39]

Definition 5. (Morishima 1974) For any $E \in \mathcal{E}$ and $\left((p,w),(\xi^{\nu})_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, agent $\nu \in \mathcal{N}$ is *exploited* if and only if $\Lambda^{\nu} > \max_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c)$ and *an exploiter* if and only if $\Lambda^{\nu} < \min_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c)$.

Definition 5 has some desirable characteristics, according to Morishima (1974, pp.616-618): the notion of exploitation is well-defined because both $\max_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c)$ and $\min_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c)$ are uniquely well-defined and positive whenever $\mathbf{0} \notin \mathcal{C}^{\nu}(p,w,\Lambda^{\nu})$;[40] and exploitation status is determined independent of price information, once the set of optimal consumption bundles is known, as in the standard Marxian approach, focusing only on production data.

According to Roemer (1981, 1982), however, Definition 5 is conceptually flawed as it identifies exploitation status (potentially) based on production techniques that will never be used by profit maximising capitalists. Like Morishima (1974), Roemer (1982) focuses on agents's optimal bundles but argues that the labour content of any bundle should be given by the minimum amount of (effective) labour necessary to produce it as net output *among profit-rate-maximising activities at given prices*, for only the latter production processes will be activated in equilibrium.

Formally, let $\mathcal{P}^{\pi}(p,w) = \left\{\alpha \in \mathcal{P} \mid \pi = \frac{p\widehat{\alpha}-w\alpha_l}{p\underline{\alpha}}\right\}$ denote the set of production processes that yield the maximum profit rate. Then, for all $c \in \mathbb{R}^n_+$, the labour content of c is:

$$l.v.(c;p,w) \equiv \min\left\{\alpha_l \mid \alpha = (-\alpha_l, -\underline{\alpha}, \overline{\alpha}) \in \phi(c) \cap \mathcal{P}^{\pi}(p,w)\right\}.$$

Then, taking account the possibility of multiple optimal consumption bundles, Roemer's (1981; 1982) definition of exploitation can be extended as follows:

Definition 6. (Roemer 1982) For any $E \in \mathcal{E}$ and $\left((p,w),(\xi^{\nu})_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, agent $\nu \in \mathcal{N}$ is *exploited* if and only if $\Lambda^{\nu} > \max_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c;p,w)$ and *an exploiter* if and only if $\Lambda^{\nu} < \min_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c;p,w)$.

While individual optimal consumption is central in Definitions 5 and 6, Roemer (1982) has also proposed an alternative approach in which agents' exploitation status is independent of their preferences over bundles of produced goods, focusing on the *maximum* and the *minimum* amounts of labour embodied in bundles that they can purchase.

[39]Clearly, Definition 5 reduces to Morishima's (1974) original formulation whenever $\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})$ is a singleton.

[40]This follows from assumptions **A0**$\sim$**A2** in Appendix A. See Roemer (1980, Proposition 2.1). The same holds for $\max_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c;p,w)$ and $\min_{c\in\mathcal{C}^{\nu}(p,w,\Lambda^{\nu})} l.v.(c;p,w)$ below.

Definition 7. (Roemer 1982) For any $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, agent $\nu \in \mathcal{N}$ is *exploited* if and only if $\Lambda^\nu > \max_{c\in\mathcal{B}(p,c^\nu)} l.v.(c)$ and *an exploiter* if and only if $\Lambda^\nu < \min_{c\in\mathcal{B}(p,c^\nu)} l.v.(c)$.

While Definition 7 measures the labour content of the relevant bundles using Morishima's price-independent method, Definition 8 focuses on profit-maximising activities.

Definition 8. (Roemer 1982) For any $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, agent $\nu \in \mathcal{N}$ is *exploited* if and only if $\Lambda^\nu > \max_{c\in\mathcal{B}(p,c^\nu)} l.v.(c;p,w)$ and *an exploiter* if and only if $\Lambda^\nu < \min_{c\in\mathcal{B}(p,c^\nu)} l.v.(c;p,w)$.

Although they preserve some important insights of standard exploitation theory (see Veneziani and Yoshihara (2012)), Definitions 5 to 8 have been criticised because exploitation status depends on counterfactual amounts of labour content. For the production activities yielding $l.v.(c^\nu)$ or $l.v.(c^\nu;p,w)$ may be different from those actually used in equilibrium. According to critics, this use of counterfactuals is theoretically undesirable and it makes exploitation an empirically vacuous notion, since the computation of $l.v.(c^\nu)$ and $l.v.(c^\nu;p,w)$ requires information that is not available, including, in Morishima's own words, "information about all the available techniques of production, actually chosen or potentially usable" (Morishima 1974, p.617).[41]

As for Definition 4, we note here that the "New Interpretation" has been criticised because, unlike Definitions 5 and 6, the actual consumption choices of the agents are only indirectly relevant to determine exploitation status, and unlike Definitions 5 and 7, the notion of exploitation depends on price information.

Next, we check that Definitions 4 to 8 are all definitions of *labour exploitation*. This will also show that axiom **LE** is indeed an appropriate domain condition as all of the main definitions in the literature satisfy it.

Consider Definition 4. For all $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$ with aggregate production activity $\alpha^{p,w}+\beta^{p,w}$, $\overline{c}^\nu = \underline{c}^\nu \equiv \tau^{c^\nu}\left(\widehat{\alpha}^{p,w}+\widehat{\beta}^{p,w}\right)$, $\alpha^{\overline{c}^\nu} = \alpha^{\underline{c}^\nu} \equiv \tau^{c^\nu}\left(\alpha^{p,w}+\beta^{p,w}\right)$, and $\alpha_l^{\overline{c}^\nu} = \alpha_l^{\underline{c}^\nu} = \tau^{c^\nu}\left(\alpha_l^{p,w}+\beta_l^{p,w}\right)$.

Consider Definition 5. For all $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu\in N}\right) \in \mathcal{RS}_E$, for all $\nu \in \mathcal{N}$:

$$\overline{c}^\nu \equiv \arg\min_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c) \quad , \quad \alpha_l^{\overline{c}^\nu} = \min_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c),$$
$$\underline{c}^\nu \equiv \arg\max_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c) \quad , \quad \alpha_l^{\underline{c}^\nu} = \max_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c).$$

Consider Definition 6. For all $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu\in N}\right) \in \mathcal{RS}_E$, for all $\nu \in \mathcal{N}$:

$$\overline{c}^\nu \equiv \arg\min_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c;p,w) \quad , \quad \alpha_l^{\overline{c}^\nu} = \min_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c;p,w),$$
$$\underline{c}^\nu \equiv \arg\max_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c;p,w) \quad , \quad \alpha_l^{\underline{c}^\nu} = \max_{c\in\mathcal{C}^\nu(p,w,\Lambda^\nu)} l.v.(c;p,w).$$

[41] For a thorough discussion, see Flaschel (1983, 2010).

Consider Definition 7. For all $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu \in N}\right) \in \mathcal{RS}_E$, for all $\nu \in \mathcal{N}$:

$$\overline{c}^\nu \equiv \arg \min_{c \in \mathcal{B}(p,c^\nu)} l.v.(c) \quad , \quad \alpha_l^{\overline{c}^\nu} = \min_{c \in \mathcal{B}(p,c^\nu)} l.v.(c),$$
$$\underline{c}^\nu \equiv \arg \max_{c \in \mathcal{B}(p,c^\nu)} l.v.(c) \quad , \quad \alpha_l^{\underline{c}^\nu} = \max_{c \in \mathcal{B}(p,c^\nu)} l.v.(c).$$

Consider Definition 8. For all $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu \in N}\right) \in \mathcal{RS}_E$, for all $\nu \in \mathcal{N}$:

$$\overline{c}^\nu \equiv \arg \min_{c \in \mathcal{B}(p,c^\nu)} l.v.(c;p,w) \quad , \quad \alpha_l^{\overline{c}^\nu} = \min_{c \in \mathcal{B}(p,c^\nu)} l.v.(c;p,w),$$
$$\underline{c}^\nu \equiv \arg \max_{c \in \mathcal{B}(p,c^\nu)} l.v.(c;p,w) \quad , \quad \alpha_l^{\underline{c}^\nu} = \max_{c \in \mathcal{B}(p,c^\nu)} l.v.(c;p,w).$$

In closing this section, we note in passing that, first, based on Flaschel's (1983; 2010) notion of *actual labour values*, another definition of exploitation can be derived that satisfies **LE**. Second, it is immediate to see that Definitions 4 to 8 all satisfy axioms **IND** and **SI**.

C Proof of Theorem 1

Theorem 1. *A definition of labour exploitation $d \in \mathcal{D}_L$ satisfies* ***RE****,* ***SI****, and* ***IND*** *if and only if $d \in \mathcal{D}_L^\tau$.*

Proof. ($\Leftarrow$): It is immediate to see that if $d \in \mathcal{D}_L^\tau$ then it satisfies **SI** and **IND**.

To see that d meets **RE**, consider $E = E\langle \mathcal{P}, \mathcal{N}, (u^\nu, s^\nu, \omega^\nu)_{\nu \in \mathcal{N}} \rangle$ and $\left((p,w),(\xi^\nu)_{\nu \in \mathcal{N}}\right) \in \mathcal{RS}_E$ with $\pi > 0$. Suppose $\mathcal{N}^{ter} \neq \varnothing$ and $\mu \in \mathcal{N}^{ter}$. Then, by construction, $\Lambda^\mu < \frac{pc^\mu}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)} (\alpha_l^{p,w} + \beta_l^{p,w})$, or $\frac{\alpha_l^\mu + \gamma^\mu}{pc^\mu} < \frac{\alpha_l^{p,w} + \beta_l^{p,w}}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)}$. As $\sum_{\nu \in \mathcal{N}} pc^\nu = p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)$ and $\sum_{\nu \in \mathcal{N}} (\alpha_l^\nu + \gamma^\nu) = \alpha_l^{p,w} + \beta_l^{p,w}$ hold, the last inequality implies that there exists $\mu' \in \mathcal{N}$ such that $\frac{\alpha_l^{\mu'} + \gamma^{\mu'}}{pc^{\mu'}} > \frac{\alpha_l^{p,w} + \beta_l^{p,w}}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)}$, which implies $\mu' \in \mathcal{N}^{ted}$.

A similar argument proves that $\mathcal{N}^{ted} \neq \varnothing$ implies $\mathcal{N}^{ter} \neq \varnothing$.

($\Rightarrow$): Consider any definition of labour exploitation $d \in \mathcal{D}_L$ satisfying **RE**, **SI**, and **IND**. Consider any $E = E\langle P, \mathcal{N}, (u^\nu, s^\nu, \omega^\nu)_{\nu \in \mathcal{N}} \rangle \in \mathcal{E}$ and any $\left((p,w),(\xi^\nu)_{\nu \in \mathcal{N}}\right) \in \mathcal{RS}_E$ such that $\pi > 0$. Let $s = \sum_{\nu \in \mathcal{N}} s^\nu$, $\omega = \sum_{\nu \in \mathcal{N}} \omega^\nu$, and $\tau^{c^\nu} = \frac{pc^\nu}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)}$ for all $\nu \in \mathcal{N}$.

By $d \in \mathcal{D}_L$, for each agent $\nu \in \mathcal{N}$, there exists a profile $\left((\overline{c}^\nu, \underline{c}^\nu), \left(\alpha^{\overline{c}^\nu}, \alpha^{\underline{c}^\nu}\right)\right)$ such that $\overline{c}^\nu, \underline{c}^\nu \in \mathcal{B}(p, c^\nu)$, $\alpha^{\overline{c}^\nu} \in \phi(\overline{c}^\nu) \cap \partial P$, $\alpha^{\underline{c}^\nu} \in \phi(\underline{c}^\nu) \cap \partial P$, and $\alpha_l^{\underline{c}^\nu} \geqq \alpha_l^{\overline{c}^\nu}$. Suppose, by way of contradiction, that $\alpha_l^{\overline{c}^\nu} = \alpha_l^{\underline{c}^\nu} = \frac{pc^\nu}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)} (\alpha_l^{p,w} + \beta_l^{p,w})$ does not hold for some $\nu \in \mathcal{N}$.

1. For the rest of the proof, starting from $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu \in \mathcal{N}}\right) \in \mathcal{RS}_E$ define the following economies and allocations.

Let $E_\tau = E\langle \mathcal{P}, \mathcal{N}, (u_\tau^\nu, s_\tau^\nu, \omega_\tau^\nu)_{\nu \in \mathcal{N}} \rangle \in \mathcal{E}(\mathcal{P}; \mathcal{N}; s; \omega)$ be such that for every $\nu \in \mathcal{N}$, $s_\tau^\nu = \tau^{c^\nu} s$, $u_\tau^\nu(c, \lambda) = pc - ws_\tau^\nu \lambda$ for any $(c, \lambda) \in \mathbb{R}_+^n \times [0,1]$, and $\omega_\tau^\nu = \tau^{c^\nu} \omega$. Let

$$(\xi_\tau^\nu)_{\nu \in \mathcal{N}} = \left(\mathbf{0}; \tau^{c^\nu} (\alpha^{p,w} + \beta^{p,w}); \tau^{c^\nu} (\alpha_l^{p,w} + \beta_l^{p,w}); c^\nu\right)_{\nu \in \mathcal{N}}.$$

Let $E_e = E\langle \mathcal{P}, \mathcal{N}, (u_e^\nu, s_e^\nu, \omega_e^\nu)_{\nu\in\mathcal{N}}\rangle \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$ such that for any $\nu \in \mathcal{N}$, $s_e^\nu = \frac{1}{N}s$, $u_e^\nu(c,\lambda) = pc - ws_e^\nu\lambda$ for any $(c,\lambda) \in \mathbb{R}_+^n \times [0,1]$, and $\omega_e^\nu = \frac{1}{N}\omega$. Let

$$(\xi_e^\nu)_{\nu\in\mathcal{N}} = \left(\mathbf{0}; \frac{1}{N}(\alpha^{p,w} + \beta^{p,w}); \frac{1}{N}(\alpha_l^{p,w} + \beta_l^{p,w}); \frac{1}{N}\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)\right)_{\nu\in\mathcal{N}}.$$

Let $c_e \equiv c_e^\nu = \frac{1}{N}\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)$.

Let $E_\theta = E\langle \mathcal{P}, \mathcal{N}, (u_\theta^\nu, s_\theta^\nu, \omega_\theta^\nu)_{\nu\in\mathcal{N}}\rangle \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$ such that $u_\theta^\nu(c,\lambda) = pc - ws_\theta^\nu\lambda$ for $(c,\lambda) \in \mathbb{R}_+^n \times [0,1]$, for all $\nu \in \mathcal{N}$ and, for some $x < \alpha_l^{p,w} + \beta_l^{p,w}$,

$$\begin{aligned}\left(s_\theta^1, (s_\theta^\nu)_{\nu\in\mathcal{N}\setminus\{1\}}\right) &= \left(x, \left(\frac{s-x}{N-1}\right)_{\nu\in\mathcal{N}\setminus\{1\}}\right); \\ \left(\omega_\theta^1, (\omega_\theta^\nu)_{\nu\in\mathcal{N}\setminus\{1\}}\right) &= \left(\theta^1\omega, (\theta^\nu\omega)_{\nu\in\mathcal{N}\setminus\{1\}}\right)\end{aligned}$$

where $\theta^1 \equiv \frac{pc_e - wx}{\pi p\omega}$ and $\theta^\nu \equiv \left(\frac{1-\theta^1}{N-1}\right)$. Let

$$\begin{aligned}\xi_\theta^1 &= \left(\mathbf{0}; \theta^1(\alpha^{p,w} + \beta^{p,w}); x; \frac{1}{N}\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)\right); \\ \xi_\theta^\nu &= \left(\mathbf{0}; \frac{1-\theta^1}{N-1}(\alpha^{p,w} + \beta^{p,w}); \frac{1}{N-1}(\alpha_l^{p,w} + \beta_l^{p,w} - x); \frac{1}{N}\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)\right)_{\nu\in\mathcal{N}\setminus\{1\}}.\end{aligned}$$

It is easy to show that if $\left((p,w), (\xi^\nu)_{\nu\in N}\right) \in \mathcal{RS}_E$ with $\pi > 0$, then $\left((p,w), (\xi_\tau^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\tau}$ with $\pi > 0$, and $\left((p,w), (\xi_e^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_e}$ with $\pi > 0$. (Observe that $\left((p,w), (\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$ implies $s \geqq (\alpha_l^{p,w} + \beta_l^{p,w})$.) Let $(\overline{c}_\tau^\nu, \underline{c}_\tau^\nu)_{\nu\in\mathcal{N}}$ and $(\overline{c}_e^\nu, \underline{c}_e^\nu)_{\nu\in\mathcal{N}}$ be, respectively, the associated ERBs. By $d \in \mathcal{D}_L$, $\left(\overline{c}_e^\nu, \underline{c}_e^\nu, \alpha_l^{\overline{c}_e^\nu}, \alpha_l^{\underline{c}_e^\nu}\right) = \left(\overline{c}_e, \underline{c}_e, \alpha_l^{\overline{c}_e}, \alpha_l^{\underline{c}_e}\right)$ for all $\nu \in \mathcal{N}$.

2. Suppose that $\alpha_l^{\overline{c}^\nu} \leqq \tau^{c^\nu}(\alpha_l^{p,w} + \beta_l^{p,w})$ for all $\nu \in \mathcal{N}$ with $\alpha_l^{\overline{c}^\mu} \leqq \alpha_l^{\underline{c}^\mu} < \tau^{c^\mu}(\alpha_l^{p,w} + \beta_l^{p,w})$ for at least some $\mu \in \mathcal{N}$. Consider E_τ and $\left((p,w), (\xi_\tau^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\tau}$ in step 1. Because $\alpha_\tau^{p,w} + \beta_\tau^{p,w} = \alpha^{p,w} + \beta^{p,w}$, $E \neq E_\tau$, and $c_\tau^\nu = c^\nu$ for all $\nu \in \mathcal{N}$, it follows from **IND** that $\left(\alpha_l^{\overline{c}_\tau^\nu}, \alpha_l^{\underline{c}_\tau^\nu}\right)_{\nu\in\mathcal{N}} = \left(\alpha_l^{\overline{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right)_{\nu\in\mathcal{N}}$ holds. Therefore, $\alpha_l^{\overline{c}_\tau^\nu} \leqq \tau^{c^\nu}(\alpha_l^{p,w} + \beta_l^{p,w}) = \Lambda_\tau^\nu$ for all $\nu \in \mathcal{N}$ with $\alpha_l^{\overline{c}_\tau^\mu} \leqq \alpha_l^{\underline{c}_\tau^\mu} < \tau^{c^\mu}(\alpha_l^{p,w} + \beta_l^{p,w}) = \Lambda_\tau^\mu$ for at least some $\mu \in \mathcal{N}$. Hence, $N_\tau^{ted} \neq \varnothing$ while $N_\tau^{ter} = \varnothing$, which contradicts **RE**.

3. A similar argument rules out the possibility that $\tau^{c^\nu}(\alpha_l^{p,w} + \beta_l^{p,w}) \leqq \alpha_l^{\underline{c}^\nu}$ for all $\nu \in \mathcal{N}$ with $\tau^{c^\mu}(\alpha_l^{p,w} + \beta_l^{p,w}) < \alpha_l^{\overline{c}^\mu} \leqq \alpha_l^{\underline{c}^\mu}$ for at least some $\mu \in \mathcal{N}$.

4. Suppose that there exist $\mu, \mu' \in \mathcal{N}$ such that $\alpha_l^{\overline{c}^\mu} > \tau^{c^\mu}(\alpha_l^{p,w} + \beta_l^{p,w})$ and $\alpha_l^{\underline{c}^{\mu'}} < \tau^{c^{\mu'}}(\alpha_l^{p,w} + \beta_l^{p,w})$. Starting from E_τ and $\left((p,w), (\xi_\tau^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\tau}$ in step 1, consider an alternative allocation $(\xi'^\nu)_{\nu\in\mathcal{N}} \equiv (\alpha'^\nu; \beta'^\nu; \gamma'^\nu; c'^\nu)_{\nu\in\mathcal{N}} = \left(\alpha_\tau^\nu; \beta_\tau^\nu; \gamma_\tau^\nu; \tau^{c^\nu}\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)\right)_{\nu\in\mathcal{N}}$. It is immediate to show that $\left((p,w), (\xi'^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\tau}$ with $\pi > 0$ and $\alpha'^{p,w} + \beta'^{p,w} = \alpha_\tau^{p,w} + \beta_\tau^{p,w}$. Let $(\overline{c}'^\nu, \underline{c}'^\nu)_{\nu\in\mathcal{N}}$ be the corresponding ERBs. By **IND**, $\left(\alpha_l^{\overline{c}'^\nu}, \alpha_l^{\underline{c}'^\nu}\right)_{\nu\in\mathcal{N}} =$

$\left(\alpha_l^{\bar{c}_\tau^\nu}, \alpha_l^{c_\tau^\nu}\right)_{\nu\in\mathcal{N}} = \left(\alpha_l^{\bar{c}^\nu}, \alpha_l^{c^\nu}\right)_{\nu\in\mathcal{N}}$, where the latter equality follows from step 2.

Next, consider E_e and $\left((p,w), (\xi_e^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_e}$ in step 1 above. Let $\chi^\nu = \frac{1}{N}\frac{1}{\tau^{c^\nu}} > 0$ for each $\nu \in \mathcal{N}$. Because $E_\tau, E_e \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$, $\left((p,w), (\xi'^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\tau}$, $\left((p,w), (\xi_e^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_e}$, $\alpha'^{p,w} + \beta'^{p,w} = \alpha^{p,w} + \beta^{p,w} = \alpha_e^{p,w} + \beta_e^{p,w}$, and $\xi_e^\nu = \chi^\nu \xi'^\nu$ for all $\nu \in \mathcal{N}$, it follows from **SI** that $\left(\alpha_l^{\bar{c}_e}, \alpha_l^{c_e}\right) = \chi^\nu\left(\alpha_l^{\bar{c}'^\nu}, \alpha_l^{c'^\nu}\right) = \chi^\nu\left(\alpha_l^{\bar{c}_\tau^\nu}, \alpha_l^{c_\tau^\nu}\right) = \chi^\nu\left(\alpha_l^{\bar{c}^\nu}, \alpha_l^{c^\nu}\right)$ holds for all $\nu \in \mathcal{N}$.

Thus, if there exist $\mu, \mu' \in \mathcal{N}$ such that $\alpha_l^{\bar{c}^\mu} > \tau^{c^\mu}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$ and $\alpha_l^{c^{\mu'}} < \tau^{c^{\mu'}}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$, then $\alpha_l^{\bar{c}_e} > \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$ and $\frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) > \alpha_l^{c_e}$, which implies $\alpha_l^{\bar{c}_e} > \alpha_l^{c_e}$, thus contradicting $d \in \mathcal{D}_L$.

5. Finally, suppose that for all $\nu \in \mathcal{N}$, $\alpha_l^{\bar{c}^\nu} \leqq \tau^{c^\nu}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) \leqq \alpha_l^{c^\nu}$ with at least one strict inequality for some $\mu \in \mathcal{N}$. Consider E_e and $\left((p,w), (\xi_e^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_e}$ in step 1 above. Using the same reasoning as in step 4, it follows that $\alpha_l^{\bar{c}_e} \leqq \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) \leqq \alpha_l^{c_e}$ with at least one strict inequality.

5.1. First, we prove first that $\pi p\omega > p\bar{c}_e - w\alpha_l^{\bar{c}_e} \geqq \frac{1}{N}\pi p\omega > 0$. The latter inequalities follow from $p\bar{c}_e - w\alpha_l^{\bar{c}_e} \geqq p\bar{c}_e - w\frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) = \frac{1}{N}\pi p\omega$ and $\pi > 0$. As for the first inequality, observe that $\pi p\omega = p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right) - w\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) \geqq \frac{1}{N}p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right) > \frac{1}{N}p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right) - w\alpha_l^{\bar{c}_e} = p\bar{c}_e - w\alpha_l^{\bar{c}_e}$. The first weak inequality holds provided N is sufficiently large given that $\pi p\omega > 0$.

5.2. Next we prove that $p\underline{c}_e - w\alpha_l^{c_e} > 0$. To see this, suppose on the contrary that $p\underline{c}_e - w\alpha_l^{c_e} \leqq 0$ holds. Consider $E_\theta \in \mathcal{E}(\mathcal{N};P;s;\omega)$ in step 1 with $x = \varepsilon \in \left(0, \alpha_l^{\bar{c}_e}\right)$. By step 5.1, $\pi p\omega > p\bar{c}_e - w\alpha_l^{\bar{c}_e} > 0$ and therefore ε can be chosen such that $1 > \theta^1 = \frac{pc_e - w\varepsilon}{\pi p\omega} > 0$.

Then, by construction, $\left((p,w), (\xi_\theta^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\theta}$ with $\pi > 0$ such that $\alpha_\theta^{p,w} + \beta_\theta^{p,w} = \alpha^{p,w} + \beta^{p,w}$ and $c_\theta^\nu = c_e$ for all $\nu \in \mathcal{N}$. Hence, noting that $E_\theta \neq E_e$, **IND** implies that $\left(\alpha_l^{\bar{c}_\theta^\nu}, \alpha_l^{c_\theta^\nu}\right)_{\nu\in\mathcal{N}} = \left(\alpha_l^{\bar{c}_e}, \alpha_l^{c_e}\right)_{\nu\in\mathcal{N}}$ holds. As $\Lambda_\theta^1 = \varepsilon < \alpha_l^{\bar{c}_e}$, we have $\{1\} \subseteq \mathcal{N}^{ter}$. In contrast, noting that $p\underline{c}_e = pc_e = \theta^\nu \pi p\omega + w\Lambda_\theta^\nu$ for all $\nu \in \mathcal{N}\setminus\{1\}$, $p\underline{c}_e - w\alpha_l^{c_e} \leqq 0$ implies that $\Lambda_\theta^\nu = \frac{1}{N-1}\left(\alpha_l^{p,w} + \beta_l^{p,w} - \varepsilon\right) < \alpha_l^{c_e}$ for all $\nu \in \mathcal{N}\setminus\{1\}$, which implies $\mathcal{N}^{ted} = \varnothing$, thus contradicting **RE**.

5.3. We consider the three cases.

Case 1: $0 \leqq \alpha_l^{c_e} - \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) < \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) - \alpha_l^{\bar{c}_e}$.

Consider $E_\theta \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$ such that $x = \alpha_l^{c_e} + \varepsilon$. By step 5.1 and $d \in \mathcal{D}_L$, $\pi p\omega > pc_e - w\alpha_l^{\bar{c}_e} \geqq pc_e - w\alpha_l^{c_e}$. By step 5.2, $pc_e - w\alpha_l^{c_e} > 0$. Therefore for a sufficiently small $\varepsilon > 0$, $1 > \theta^1 = \frac{pc_e - w\left(\alpha_l^{c_e} + \varepsilon\right)}{\pi p\omega} > 0$. Then, by construction, $\left((p,w), (\xi_\theta^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\theta}$ with $\pi > 0$ such that $\alpha_\theta^{p,w} + \beta_\theta^{p,w} = \alpha^{p,w} + \beta^{p,w}$ and $c_\theta^\nu = c_e$ for all $\nu \in \mathcal{N}$. Hence, noting that $E_\theta \neq E_e$, **IND** implies that $\left(\alpha_l^{\bar{c}_\theta^\nu}, \alpha_l^{c_\theta^\nu}\right)_{\nu\in\mathcal{N}} = \left(\alpha_l^{\bar{c}_e}, \alpha_l^{c_e}\right)_{\nu\in\mathcal{N}}$ holds.

As $\Lambda_\theta^1 > \alpha_l^{c_e}$, we have $\{1\} \subseteq \mathcal{N}^{ted}$. In contrast, $\alpha_l^{c_e} - \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) < \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) - \alpha_l^{\bar{c}_e}$ ensures that $\alpha_l^{\bar{c}_e} \leqq \Lambda_\theta^\nu = \frac{1}{N-1}\left(\alpha_l^{p,w} + \beta_l^{p,w} - \alpha_l^{c_e} - \varepsilon\right)$ for all $\nu \in \mathcal{N}\setminus\{1\}$, which implies $\mathcal{N}^{ter} = \varnothing$, thus contradicting **RE**.

Case 2: $\alpha_l^{c_e} - \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) > \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) - \alpha_l^{\bar{c}_e} \geqq 0$.

Consider $E_\theta \in \mathcal{E}(\mathcal{P};\mathcal{N};P;s;\omega)$ such that $x = \alpha_l^{\bar{c}_e} - \varepsilon$. By step 5.1, $\pi p\omega > p\bar{c}_e - w\alpha_l^{\bar{c}_e} \geqq \frac{1}{N}\pi p\omega > 0$. Therefore for a sufficiently small $\varepsilon > 0$, $1 > \theta^1 = \frac{pc_e - w\left(\alpha_l^{\bar{c}_e} - \varepsilon\right)}{\pi p\omega} > 0$. By construction, $\left((p,w),(\xi_\theta^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\theta}$ with $\pi > 0$ such that $\alpha_\theta^{p,w} + \beta_\theta^{p,w} = \alpha^{p,w} + \beta^{p,w}$ and $c_\theta^\nu = c_e$ for all $\nu \in \mathcal{N}$. Hence, noting that $E_\theta \neq E_e$, **IND** implies that $\left(\alpha_l^{\bar{c}_\theta^\nu}, \alpha_l^{\underline{c}_\theta^\nu}\right)_{\nu\in\mathcal{N}} = \left(\alpha_l^{\bar{c}_e}, \alpha_l^{\underline{c}_e}\right)_{\nu\in\mathcal{N}}$ holds.

As $\Lambda_\theta^1 < \alpha_l^{\bar{c}_e}$, we have $\{1\} \subseteq \mathcal{N}^{ter}$. In contrast, $\alpha_l^{\underline{c}_e} - \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) > \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) - \alpha_l^{\bar{c}_e}$ ensures that $\alpha_l^{\underline{c}_e} \geqq \Lambda_\theta^\nu = \frac{1}{N-1}\left(\alpha_l^{p,w} + \beta_l^{p,w} - \alpha_l^{\bar{c}_e} + \varepsilon\right)$ for all $\nu \in \mathcal{N}\setminus\{1\}$, which implies $\mathcal{N}^{ted} = \varnothing$ thus contradicting **RE**.

Case 3: $\alpha_l^{\underline{c}_e} - \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) = \frac{1}{N}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right) - \alpha_l^{\bar{c}_e}$.

We can assume that $N > 2$ without loss of generality. Then, using the same construction as in case 2, it is possible to show that $E_\theta \in \mathcal{E}(\mathcal{P};\mathcal{N};s;\omega)$ and $\left((p,w),(\xi_\theta^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_{E_\theta}$ with $\pi > 0$ such that $\mathcal{N}^{ter} = \{1\}$ and $\mathcal{N}^{ted} = \varnothing$, thus contradicting **RE**.

In summary, for any definition of labour exploitation $d \in \mathcal{D}_L$ satisfying **RE**, **SI**, and **IND**, it must be $d \in \mathcal{D}_L^r$: for any $E \in \mathcal{E}$ and any $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$ with $\pi > 0$, $\alpha_l^{\bar{c}^\nu} = \alpha_l^{\underline{c}^\nu} = \frac{pc^\nu}{p\left(\widehat{\alpha}^{p,w} + \widehat{\beta}^{p,w}\right)}\left(\alpha_l^{p,w} + \beta_l^{p,w}\right)$ for all $\nu \in \mathcal{N}$. □

D Replication Invariance

Theorem 1 is derived under the assumption that the number of agents is sufficiently large. In this appendix we present an axiom that allows to derive the same characterisation for any $N \geqq 2$. The axiom captures an invariance property that a definition of exploitation should satisfy when population varies. Let $E = E\langle\mathcal{P},\mathcal{N},(u^\nu,s^\nu,\omega^\nu)_{\nu\in\mathcal{N}}\rangle \in \mathcal{E}$. For any $k \in \mathbb{N}$, the *k-replica of E* is defined as the economy $E_k = E\langle\mathcal{P},\mathcal{N}_k,(u^\nu,s^\nu,\omega^\nu)_{\nu\in\mathcal{N}_k}\rangle \in \mathcal{E}$ such that $\mathcal{N}_k = \{1,\ldots,kN\}$ and for all $\mu \in \mathcal{N}_k$ and $\nu \in \mathcal{N}$, $(u^\mu,s^\mu,\omega^\mu) = (u^\nu,s^\nu,\omega^\nu)$ whenever $\mu = (i-1)N + \nu$ for some $i = 1,...,k$.

If $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, then for any $k \in \mathbb{N}$, the *k-replica RS for* E_k is denoted as $\left((p,w),(\xi^\mu)_{\mu\in\mathcal{N}_k}\right) \in \mathcal{RS}_{E_k}$, where for all $\mu \in \mathcal{N}_k$ and $\nu \in \mathcal{N}$, $\xi^\mu = \xi^\nu$ whenever $\mu = (i-1)N + \nu$ for some $i = 1,...,k$.

Let $\left(\bar{c}^\nu, \underline{c}^\nu; \alpha_l^{\bar{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right)_{\nu\in\mathcal{N}}$ and $\left(\bar{c}^\mu, \underline{c}^\mu; \alpha_l^{\bar{c}^\mu}, \alpha_l^{\underline{c}^\mu}\right)_{\nu\in\mathcal{N}_k}$ be the ERBs and labour contents associated, respectively, with $E \in \mathcal{E}$ and $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$ and with E_k and $\left((p,w),(\xi^\mu)_{\mu\in\mathcal{N}_k}\right) \in RS_{E_k}$. Then:

Replication Invariance (RI): *For all* $E \in \mathcal{E}$ *and* $\left((p,w),(\xi^\nu)_{\nu\in\mathcal{N}}\right) \in \mathcal{RS}_E$, *and all* E_k *and* $\left((p,w),(\xi^\mu)_{\mu\in\mathcal{N}_k}\right) \in \mathcal{RS}_{E_k}$: *for all* $\mu \in \mathcal{N}_k$ and $\nu \in \mathcal{N}$, $\left(\alpha_l^{\bar{c}^\mu}, \alpha_l^{\underline{c}^\mu}\right) = \left(\alpha_l^{\bar{c}^\nu}, \alpha_l^{\underline{c}^\nu}\right)$ *whenever* $\mu = (i-1)N + \nu$ *for some* $i = 1,...,k$.

It is immediate to see that Definitions 4 to 8 all satisfy **RI**.

E Measuring exploitation: Earned income share approach

E.1 Methodology and assumptions

Assume that all differences in wages reflect differences in productivity and all individuals are paid the same wage per unit of effective labour, as in a neoclassical perfectly competitive economy. Under this assumption, the observed hourly wage is $w^\nu = \bar{w}s^\nu$, where $\bar{w}$ is the (constant) wage per unit of effective labour, and effective labour can measured as $\Lambda^\nu = \frac{w^\nu \lambda^\nu}{\bar{w}}$, where $\bar{w}$ can be ignored since it is a constant common to all individuals.

This implies that exploitation only arises from differences across individuals in the share of unearned income in total income. Indeed exploitation intensity for individual ν can be written as follows

$$\ln\left(\frac{w^\nu \lambda^\nu}{I^\nu}\right) - \ln\left(\frac{w\lambda}{I}\right),$$

where w, λ and I are respectively the average hourly wage, total hours worked, and total income in the economy.

In computing this measure using CPS data, we consider the business and/or professional practice income earned by self-employed persons as equivalent to wage income. Therefore $w\lambda$ corresponds to total *earned* income (as opposed to capital incomes).[42]

E.2 Results

Results are displayed in Figures E.1 to E.3 and Table E.1.

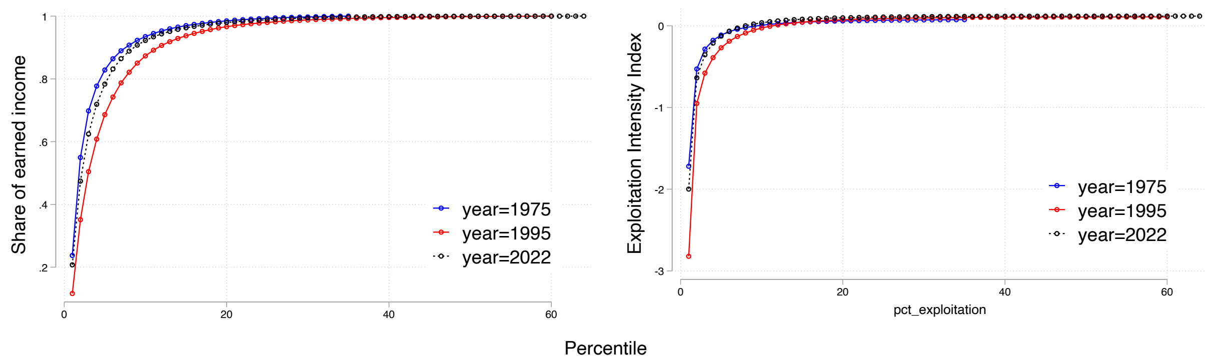
The share of earned income in total market income (the measure of labour-income ratio according to this approach) is one or close to one for around 80% of the sample, and above 0.8 for 95% of the sample (left panel of Figure E.1).

As a result, most of the population appears to be modestly exploited, while the small share of the population with large enough wealth to receive substantial unearned income has a large negative exploitation index (right panel of Figure E.1 and Figure E.2). The Gini index of the earned income share (and therefore for the exploitation intensity in this approach) is higher and it increased more than the Gini index of income (Figure E.3).

Unsurprisingly, in this approach, receiving unearned income ('rentier' variable) or being retired is associated with lower exploitation intensity, while business ownership is not associated with lower exploitation – given that high earnings from own business are entirely attributed to productivity. Moreover, given that this approach implicitly attributes racial and gender wage gaps to differences in productivity, the racial and gender gaps in exploitation intensity are strongly attenuated. The racial exploitation gap is much smaller, and the gender exploitation gap is even reversed, when using this approach (Table E.1).

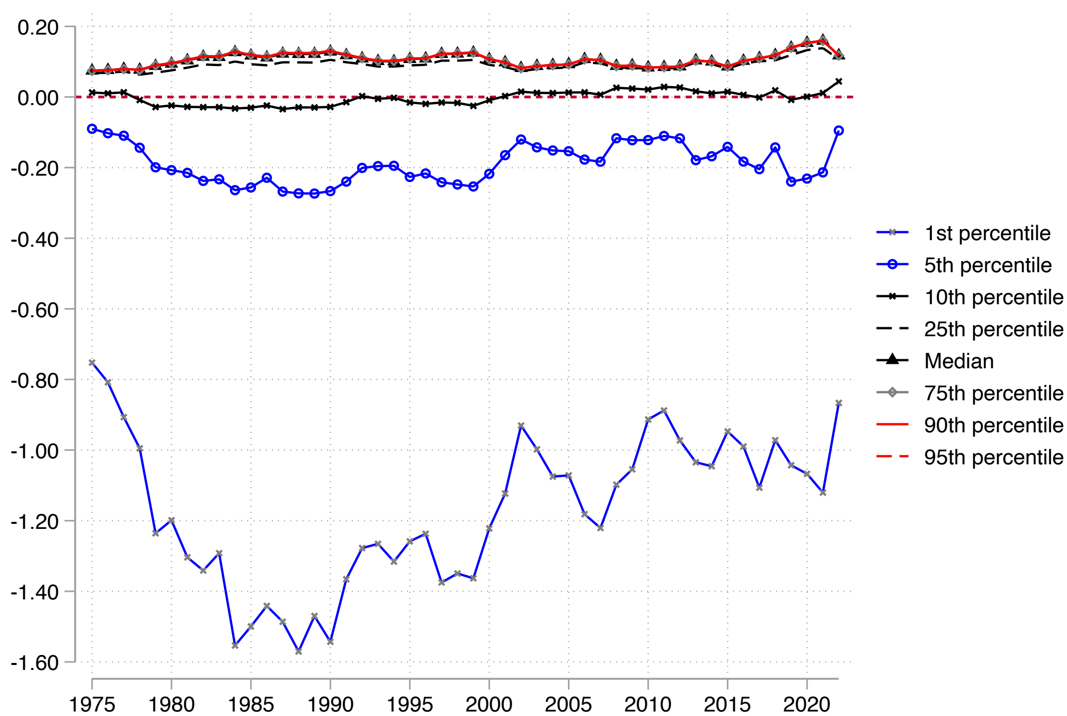
[42] As in the main analysis, we remove from the sample individuals with zero or negative market income and those who do not work because they are retired, homemakers, or students. For observations outside these categories with zero or negative earned income, and who receive non-negligle market income (defined as at least 7,000 1999 dollars as in the main analysis), we attribute 1 1999 dollar of earned income, so that their exploitation index can be computed (consistent with how we treat observations with 0 hours worked in the main analysis).

Figure E.1: Earned income share and resulting exploitation intensity index by percentile in selected years



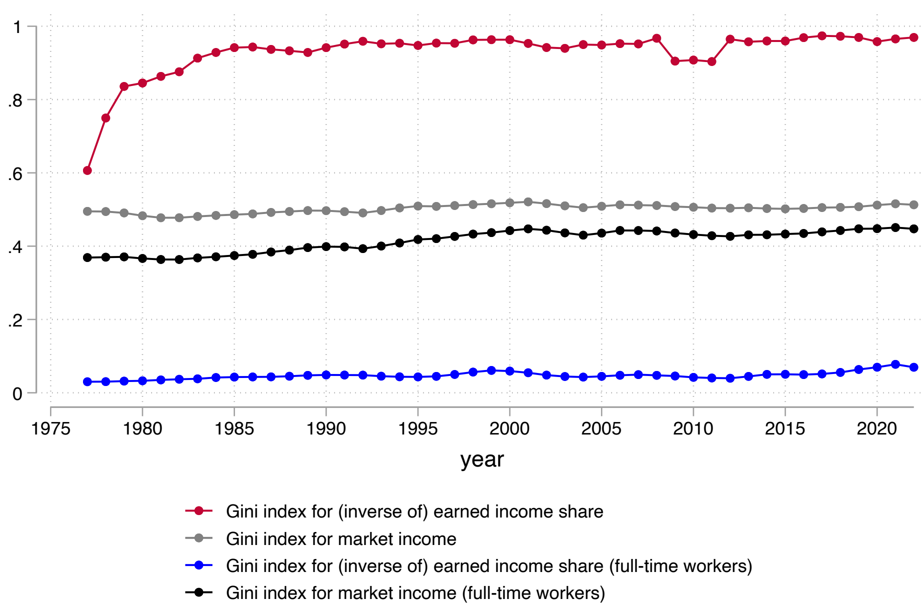
Notes: Estimated from CPS data. Exploitation intensity equals the logarithmic deviation of the individual effective labour-income ratio from the economy-wide ratio (see Equation (6)*). Effective labour measured through earned income. See main text and Appendix E for details.*

Figure E.2: Exploitation intensity index by percentile ('earned income share' approach)



Notes: Estimated from CPS data. Exploitation intensity equals the logarithmic deviation of the individual effective labour-income ratio from the economy-wide ratio (see Equation (6)*). Effective labour measured through earned income. See main text and Appendix E for details.*

Figure E.3: Inequality: Gini index for earned income share



Notes: Estimated from CPS data. Series smoothed with a 3-years backward moving average. Full-time workers are individuals who worked full-time (≥ 35 hours per week) for the full year. See main text for sample and definitions.

Table E.1: OLS estimates for the relation between exploitation and personal characteristics ('earned income share' approach)

	(1) Full Sample	(2) 1975-1989	(3) 1990-2004	(4) 2005-2022
female	-0.03***	-0.05***	-0.03***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
rentier	-0.10***	-0.11***	-0.11***	-0.09***
	(0.00)	(0.00)	(0.00)	(0.00)
entrepreneur	0.01***	0.02***	0.02***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
retired	-0.13***	-0.13***	-0.11***	-0.14***
	(0.00)	(0.01)	(0.01)	(0.00)
unemployed	0.00	0.01***	0.00***	-0.00**
	(0.00)	(0.00)	(0.00)	(0.00)
black	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
N	3995466	1120263	1234893	1640310

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is the exploitation intensity index. Estimated from CPS data. All specifications control for a full set of age-by-year fixed effects.

F Measuring exploitation: Mincer-equation approach

F.1 Methodology and assumptions

Assume $w^\nu = s^\nu \psi^\nu$, where w^ν is ν's hourly wage, s^ν is the hourly contribution of individual ν's skills in production (as defined in main text), and ψ^ν is a multiplicative error term that captures other factors influencing ν's wage. We (crucially) assume that these other factors are independent of skills: $E(\psi^\nu | s^\nu) = E(\psi^\nu) = 1$.

Further assume that productive skills are fully determined by a vector of observable variables, denoted as $\boldsymbol{x}$: $s^\nu = f(\boldsymbol{x}^\nu)$. Specifically, $\boldsymbol{x}$ includes years of formal education and job experience: $s^\nu = f(educ^\nu, exper^\nu)$. Assume the following functional relation:

$$s^\nu = \exp\left[\beta_0 + \beta_1 educ^\nu + \beta_2 exper^\nu + \beta_3 (exper^\nu)^2\right].$$

Effective labour performed is then equal to

$$\Lambda^\nu = s^\nu \lambda^\nu = \exp\left[\beta_0 + \beta_1 educ^\nu + \beta_2 exper^\nu + \beta_3 (exper^\nu)^2\right] \lambda^\nu = E(w^\nu | \boldsymbol{x}) \lambda^\nu,$$

where $E(w^\nu | \boldsymbol{x})$ is the conditional expected value of the wage based on education and experience and λ^ν is hours worked.

We can then write exploitation intensity for an individual ν as follows:

$$\ln\left(\frac{E(w^\nu | \boldsymbol{x}) \lambda^\nu}{I^\nu}\right) - \ln\left(\frac{E(w)\lambda}{I}\right),$$

where $E(w)$ is the mean wage in the population, λ is aggregate hours worked and I is aggregate income. Given our assumptions, $E(w^\nu | \boldsymbol{x})$ can be estimated from the OLS coefficients $\hat{\beta}_s$ of a 'Mincer' regression (Heckman, Lochner, and Todd 2006), in which the log of the wage is regressed on years of formal education, experience, and experience squared.[43] In implementing this approach, we allow all β coefficients to vary by year, in order to allow for changes over time in the returns to education.

F.2 Results

Figures F.1 to F.7 and Table F.1 present the results of this 'Mincer-equation' approach to estimating exploitation.

Exploitation intensity for the most exploited percentiles is by construction lower compared to the approach in the main text, but dynamics are similar (Figures F.3 to F.5). The most exploited half of the population has experienced a substantial increase in exploitation intensity over the sample period, while the 10 percent least exploited (or exploiters) have seen a further reduction in their negative exploitation rate. The magnitude of these changes is comparable to that reported in the main text using simple labour-income ratios.

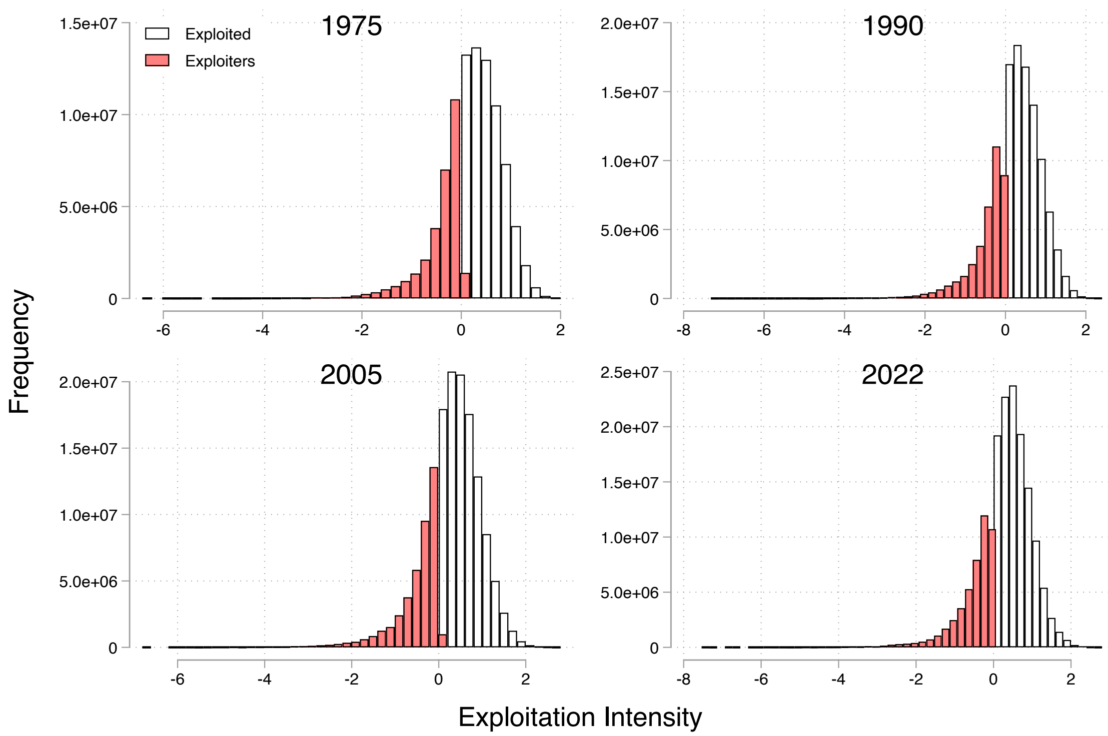
Indicators of inequality in exploitation (Figures F.6 and F.7) have increased over time. Both the standard deviation of the exploitation index and its Gini index have increased

[43] Formally: $\ln(w^\nu) = \beta_0 + \beta_1 educ^\nu + \beta_2 exper^\nu + \beta_3 (exper^\nu)^2 + \epsilon^\nu$, where $\epsilon = ln(\psi)$.

substantially more than those computed from simple labour-income ratios, and more than indicators of income inequality.

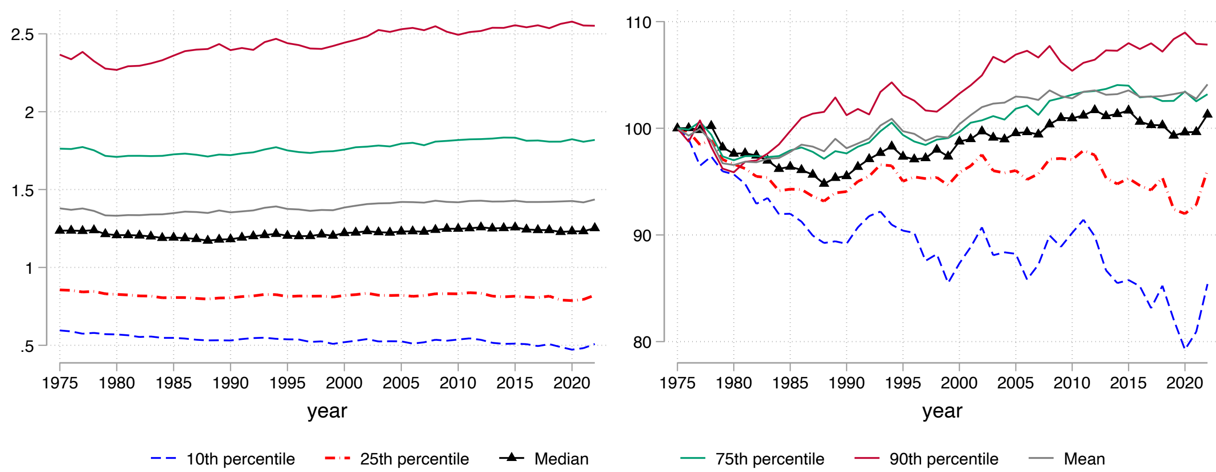
Results from regressions of exploitation intensity on gender, racial and socioeconomic characteristics produce results very similar to those in the main text (Table F.1).

Figure F.1: Distribution of Exploitation Intensity in selected years (Mincer-equation approach)



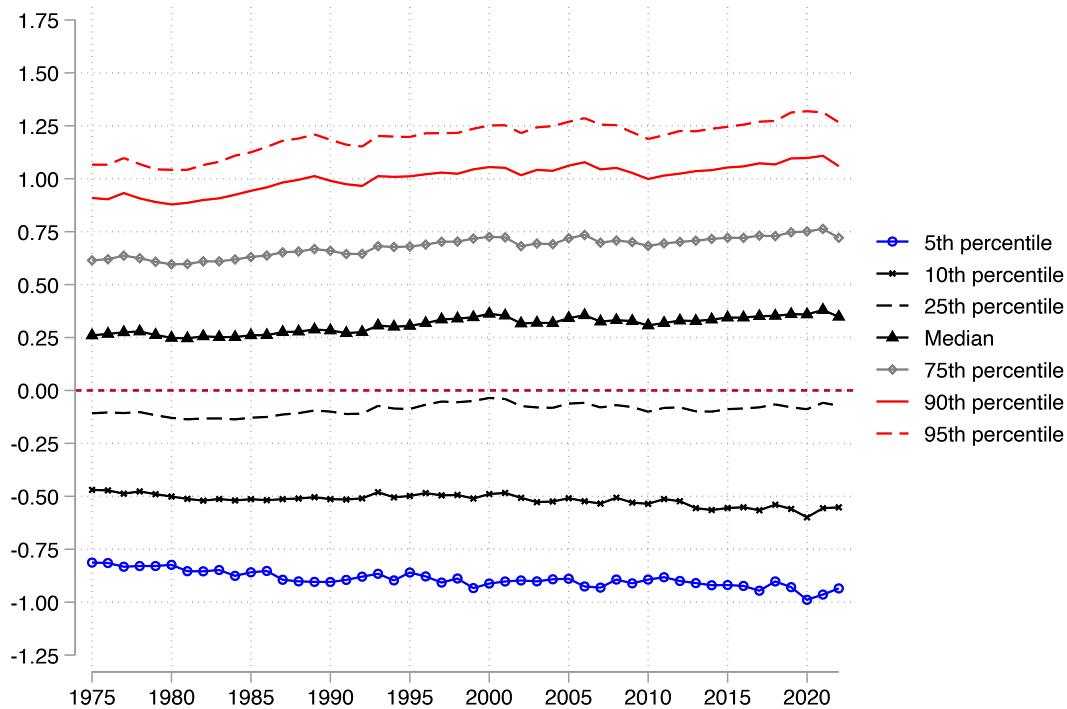
Notes: Estimated from ASEC CPS data. Exploitation intensity equals the logarithmic deviation of the individual effective labour-income ratio from the economy-wide ratio (see Equation (6)*). Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. 'Exploiters' are observations with a individual labour-income ratio below the aggregate ratio. See main text for detailed definitions.*

Figure F.2: labour-income ratios by percentile (constant 1999 dollars) (Mincer-equation approach)



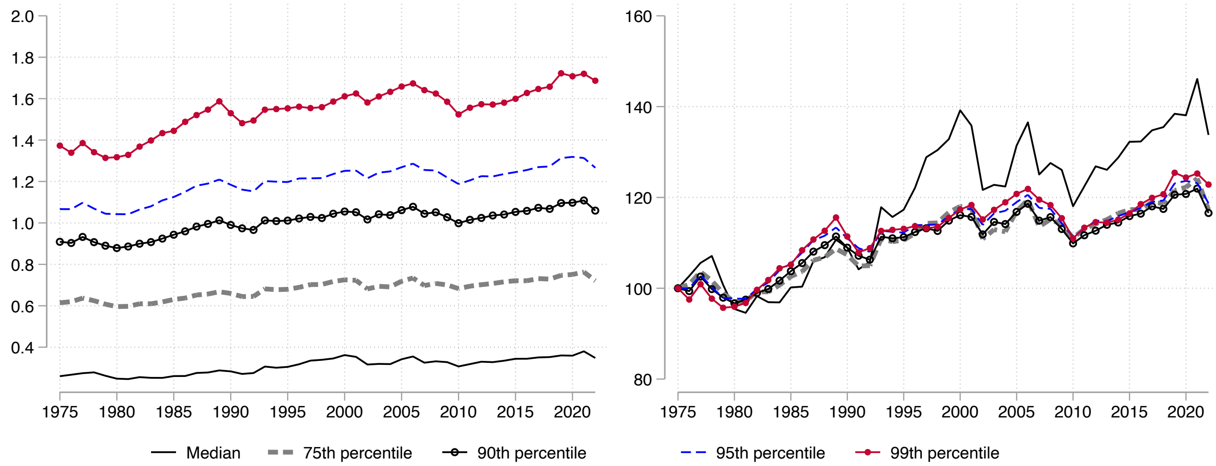
Notes: Effective labour-income ratios are defined as hours worked during the year divided by total market income (in 1999 constant USD). Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. The left panel displays levels (hours worked per dollar). In the right panel, 1975 values are normalized to 100 to highlight dynamics. See main text for sample and detailed definitions.

Figure F.3: Exploitation intensity index by percentile (Mincer-equation approach)



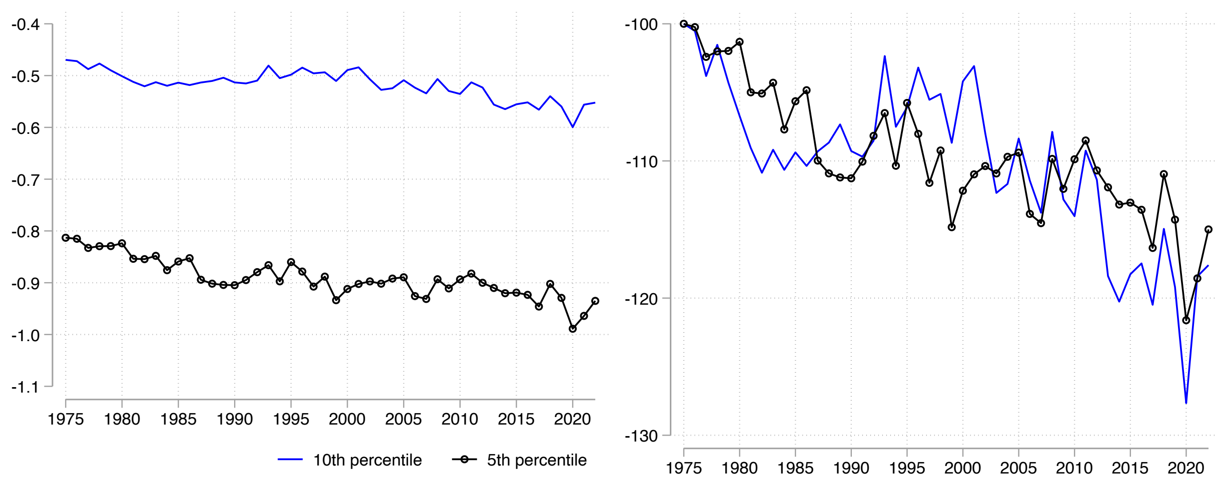
Notes: Exploitation intensity equals the logarithmic deviation of the individual effective labour-income ratio from the economy-wide ratio (see Equation (6)*). Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. See main text for sample and detailed definitions.*

Figure F.4: Exploitation intensity for the 50% most exploited (Mincer-equation approach)



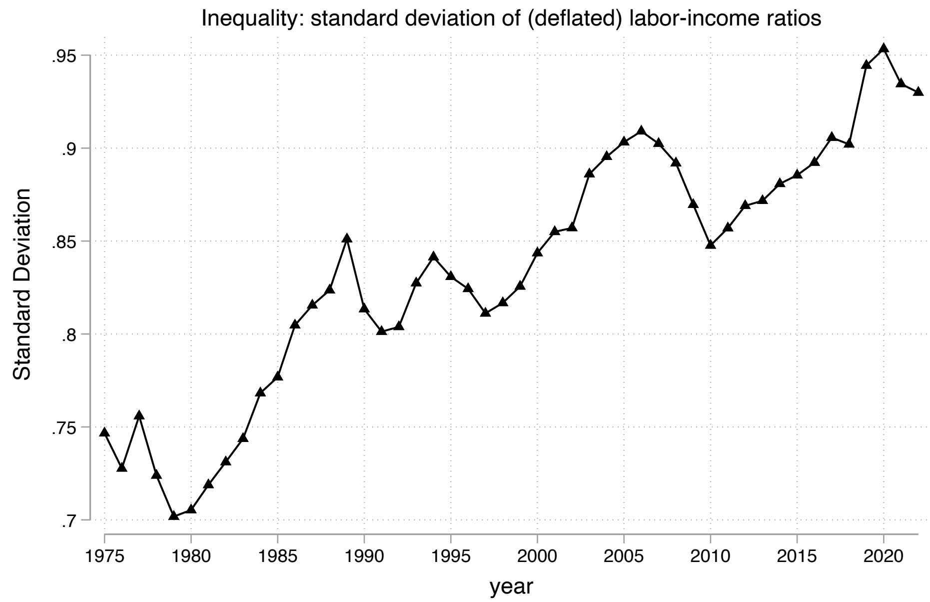
Notes: Exploitation intensity equals the logarithmic deviation of the individual effective labour-income ratio from the economy-wide ratio (see Equation (6)*). Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. The left panels displays rates of exploitation. In the right panel, 1975 rates are normalized to 100 to highlight dynamics. See main text for sample and detailed definitions.*

Figure F.5: Exploitation intensity for the 10% least exploited (aka exploiters) (Mincer-equation approach)



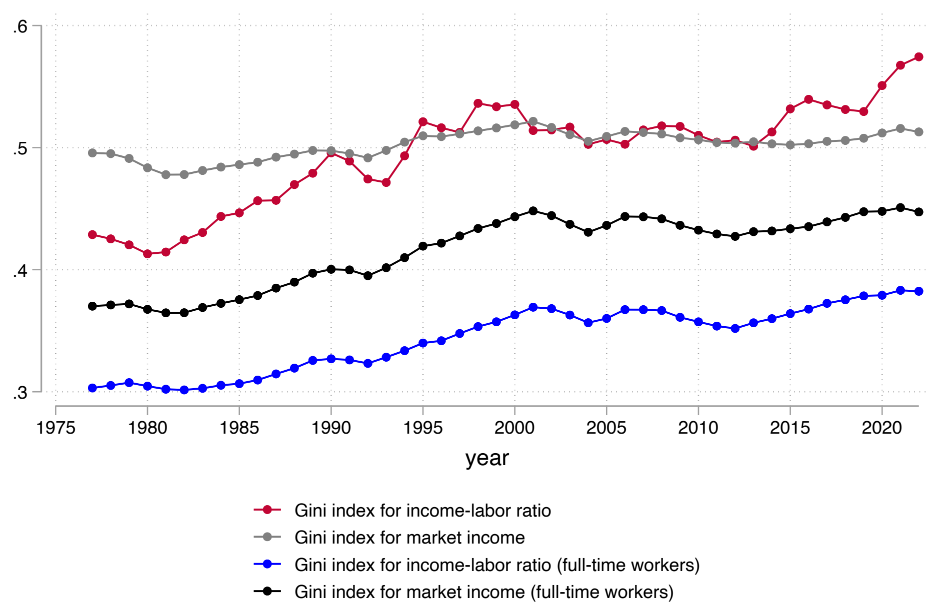
Notes: Exploitation intensity equals the logarithmic deviation of the individual effective labour-income ratio from the economy-wide ratio (see Equation (6)*). Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. The left panels displays rates of exploitation. In the right panel, 1975 rates are normalized to -100 to highlight dynamics. See main text for sample and detailed definitions.*

Figure F.6: Inequality: Standard deviation of effective labour-income ratios (1999 constant US dollars) (Mincer-equation approach)



Notes: Standard deviation of effective labour-income ratio, measured in constant 1999 dollars. Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. See main text for sample and definitions.

Figure F.7: Inequality: Gini index (Mincer-equation approach)



Notes: Series smoothed with a 3-years backward moving average. Effective labour measured through a Mincer-equation approach, as described in Section 5.1 and Appendix F. Full-time workers are individuals who worked full-time (≥ 35 hours per week) for the full year. See main text for sample and definitions.

Table F.1: OLS estimates for the relation between exploitation and personal characteristics ('Mincer-equation' approach)

	(1) Full Sample	(2) 1975-1989	(3) 1990-2004	(4) 2005-2022
female	0.24***	0.30***	0.23***	0.22***
	(0.00)	(0.00)	(0.00)	(0.00)
rentier	-0.30***	-0.27***	-0.31***	-0.31***
	(0.00)	(0.00)	(0.00)	(0.00)
entrepreneur	-0.62***	-0.53***	-0.62***	-0.68***
	(0.00)	(0.00)	(0.00)	(0.00)
retired	-0.07***	-0.06***	-0.04***	-0.09***
	(0.00)	(0.01)	(0.01)	(0.01)
unemployed	0.00	0.03***	0.00	-0.00
	(0.00)	(0.01)	(0.00)	(0.00)
black	0.08***	0.08***	0.07***	0.09***
	(0.00)	(0.00)	(0.00)	(0.00)
N	4022936	1130380	1245347	1647209

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Dependent variable is the exploitation intensity index, estimated using the 'Mincer-equation' approach. Estimated from CPS data. All specifications control for a full set of age-by-year fixed effects.

G Construction of the exploitation index from the General Social Survey (GSS)

We compute our exploitation measures also in the General Social Survey (GSS), a nationally representative survey carried out in the US since 1972. For our purposes, the available GSS sample covers the 1973-2017 period.[44] The GSS has the advantage of including questions about well-being, class identification and political orientation, thus allowing an individual-level analysis of the correlation between exploitation intensity and these outcomes. However, it provides much smaller yearly samples relative to the CPS.

Further, it allows a much rougher estimation of individual labour-income ratios. As for income, respondents are asked about the bracket in which their pre-tax income falls. We impute categorical mid-points to observations to obtain a continuous income variable. Personal income is topcoded at a relatively low level (170,000 USD at current prices in the 2016-2018 editions) and, unlike in the CPS, there is no procedure for reconstructing the distribution above the threshold. The GSS only provides data on total pre-tax income, so we are not able to subtract transfers to properly compute market income.

As for labour, respondents provide information about the hours the worked during the week prior to the survey. We therefore use 'hours worked last week' as a proxy for the average number of hours worked per week last year, and assume that every respondent who is in the labour force works 48 weeks per year.[45]

We use income and labour thus defined to compute an approximate measure of the individual exploitation intensity index. As for the CPS sample, we remove observations with a labour-income ratio that implies hourly earnings lower than two-thirds of the prevailing minimum wage.

The resulting sample includes 29,717 observations with a well-defined exploitation intensity index during 1973-2017. The distribution of labour-income ratios and exploitation intensity from the GSS samples is shown in Figures G.1 and G.2.

The socio-economic outcomes analysed in Section 7 are obtained as follows:

- 'Happy' is a binary variable equal to 1 if the respondent declares to be 'very happy' and 0 otherwise (the other two options are 'pretty happy' or 'not too happy').

- 'Healthy' is a binary variable equal to 1 if the respondent declares to be in 'excellent' or 'good' health, and 0 otherwise (the other two options are 'fair' or 'poor').

- 'Job satisfied' is a binary variable equal to 1 if the respondent declares to be 'very

[44] The GSS is carried out by the National Opinion Research Center at the University of Chicago and was retrieved from its official website (Smith, Davern, Freese, and Morgan 2021). The GSS asks respondents about income in the previous year, therefore the 1973-2017 period corresponds to the 1974-2018 surveys. Before 1974, the GSS does not report personal income. The GSS is not administered every year. In our sample period, the GSS was carried out every year from 1974 to 1994 (except for 1979, 1981, and 1992) and every other year since 1994. GSS surveys carried out after 2018 (in 2020 and 2022) are not comparable to the previous ones because of changes in survey procedures after the COVID-19 pandemics. For this reason we do not include them in our analysis.

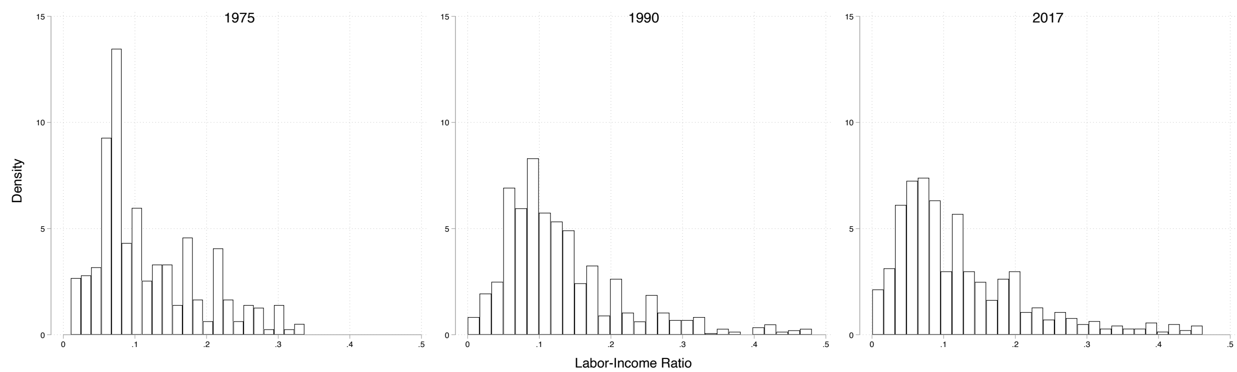
[45] Thus, for a GSS survey conducted in year $t+1$, we approximate the number of hours worked in year t Λ_t as hours worked last week times forty eight.

satisfied' or 'moderately satisfied' with their job, 0 otherwise (the other two options are 'a little dissatisfied' and 'very dissatisfied').

- 'Working Class' is a binary variable equal to 1 if the respondent declares to identify with 'lower class' or 'working class', and 0 otherwise (other two options are 'middle class' and 'upper class').

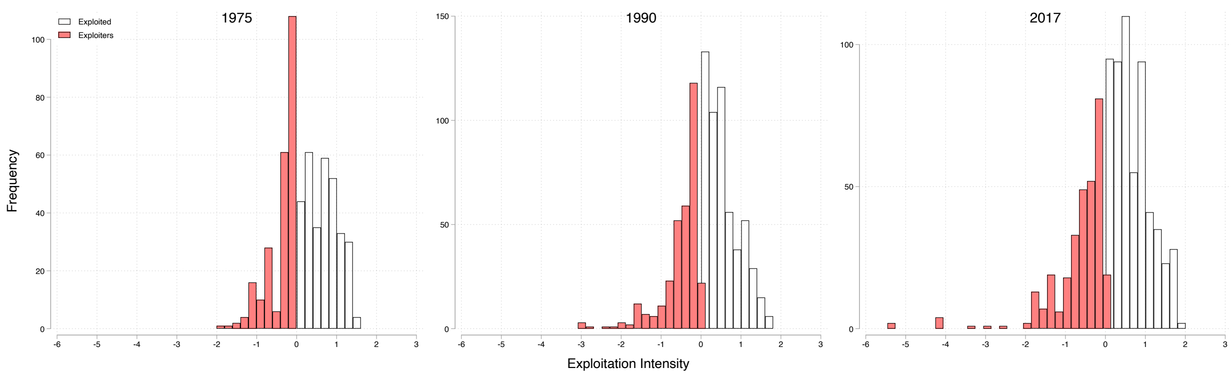
- 'Democrat' is a binary variable equal to 1 if the respondent's declared political affiliation is 'strong democrat', 'not very strong democrat' or 'independent, close to democrat', and 0 if it is 'independent (neither, no response)', 'independent, close to republican', 'not very strong republican', 'strong republican'. People identifying with 'other party' were excluded from this analysis (coded as missing values for this variable) since these parties could be rightwing or leftwing and we have no way to distinguish.

Figure G.1: Distribution of (deflated) labour-income ratios in selected years in the GSS sample



Notes: Estimated from GSS data. Labour-income ratios are defined as hours worked during the year divided by total pre-tax income (in 1986 constant USD). Therefore the unit of measure is hours worked per dollar. See main text for sample and detailed definitions.

Figure G.2: Distribution of exploitation intensity in selected years in the GSS sample



Notes: Estimated from GSS data. Exploitation intensity equals the logarithmic deviation of the individual labour-income ratio from the economy-wide ratio (see Equation (6)*). See main text for sample and detailed definitions.*