CQE Center for Quantitative Economics

Inequality Measurement – A unifying framework for dynamics, multidimensionality, and uncertainty

Simon Haastert [†], Christian Schluter [#], Mark Trede [†]

111/2025

[†] Department of Economics, University of Münster, Germany
[#] Aix Marseille School of Economics, Aix-Marseille Université, France and Department of Economics, University of Southampton, United Kingdom

wissen.leben
Universität Münster

Inequality Measurement – A unifying framework for dynamics, multidimensionality, and uncertainty*

Simon Haastert† Christian Schluter‡ Mark Trede §

January 31, 2025

Abstract

We propose a new unifying framework for the measurement of economic inequality that incorporates dynamics, multidimensionality, and uncertainty. A dynamic social evaluation function aggregates the individual value functions, which are the solutions to the individual dynamic programming problems. The proposed inequality measure is a dynamic and multivariate version of the well-known Atkinson index. For each state variable, we minimize the total resources necessary to keep current social welfare constant. Shadow values corresponding to these optimization problems capture the trade-off across dimensions in a common metric and define the weights of each state dimension for the multivariate inequality index. This linearly decomposable index measures the weighted minimum share of the total amount of the state variables necessary to attain the current level of social welfare.

The construction of this new inequality index is illustrated in a structural household life-cycle model with child care which is estimated on PSID data for married couples in the US in 2000. State variables include household wealth, and stochastic male and female wages. We show how the states do not exclusively determine the value of the inequality measure but also the future actions that arise from them, leading to important differences between the dynamic and static inequality measures.

Finally, we demonstrate how a reduced-form approach can be used instead of a structural model. We deploy offline reinforcement learning techniques, specifically offline Monte Carlo prediction, to estimate the value functions, and show that in this empirical setting, reduced-form and structural approaches yield almost identical dynamic inequality measures.

*Acknowledgements: All authors are grateful for financial support from the ANR-DFG (grant ANR-19-FRAL-0006-01). Computations were carried out on the HPC cluster PALMA II of the University of Münster, subsidised by the DFG (grant INST 211/667-1). Schluter also acknowledges core institutional funding from grant ANR-17-EURE-0020.

†Universität Münster, Am Stadtgraben 9, D-48143, Münster, Germany, (simon.haastert@wiwi.uni-muenster.de)

‡Aix-Marseille Université (Aix Marseille School of Economics), CNRS & EHESS, 5 Boulevard Maurice Bourdet CS 50498, 13205 Marseille Cedex 01, France, and Department of Economics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK (christian.schluter@univ-amu.fr and https://christianschluter.github.io/)

§Universität Münster, Am Stadtgraben 9, D-48143, Münster, Germany, (mark.trede@uni-muenster.de)

Keywords multidimensional inequality; social welfare function; structural econometrics; dynamic optimization; reinforcement learning; deep learning

JEL Classification D31 [Personal Income, Wealth, and Their Distributions], D63 [Equity, Justice, Inequality, and Other Normative Criteria and Measurement], D15 [Intertemporal Household Choice; Life Cycle Models and Saving], C61 [Optimization Techniques; Programming Models; Dynamic Analysis], I32 [Measurement and Analysis of Poverty]

1. Introduction

As economic inequality has risen steadily within and between most countries in the last decades, so has public and policy concern about these increasing disparities. Empirical investigations usually deploy inequality indices, such as the well-known Gini coefficient, to quantify a static and cross-sectional, distributional snapshot of a univariate outcome such as income. Yet, incomes evolve over time, individuals differ in terms of age, wealth, or needs, and are exposed to income risk. The inequality literature addresses these challenges usually one at a time rather than comprehensively.[1] By contrast, we propose a new unifying framework that incorporates dynamics, multidimensionality and uncertainty.

To this end, we adopt a welfarist perspective in the tradition of Atkinson (1970) whose point of departure is a social evaluation function that aggregates static individual utility functions. Instead, we consider a *dynamic social evaluation function* that aggregates the individual value functions, which are the solutions to the individual dynamic optimization problems. Our setting is inherently multidimensional since the individual value function may be a function of multiple state variables, which the dynamic social evaluation function inherits. For each state variable, we minimize the total resources necessary to keep social welfare constant. The shadow values (Lagrangian multipliers) that correspond to these optimization problems capture the trade-off across dimensions in a common metric. We demonstrate that these shadow values can be utilized as weights of each state dimension to arrive at a multivariate inequality index. This index measures the weighted minimum share of the total amount of state variables that is necessary to reach the current level of social welfare. Further, we show that the index of total inequality decomposes linearly as the weighted sum of each state variable's univariate inequality measure.

In the second part of the paper, we illustrate the construction of our dynamic and multidimensional inequality measure in a specific setting. In particular, we consider the household life-cycle model with child care presented in Blundell et al. (2018), which is then taken to US panel data for married couples from the Panel Study of Income Dynamics (PSID) in 2000.

[1] The *Handbook of Income Distribution* illustrates this separation, see e.g. the survey chapters on inequality measurement (Cowell 2000), income mobility (Jäntti and Jenkins 2015), equality of opportunity (Roemer and Trannoy 2015), and multidimensionality (Aaberge and Brandolini 2015).

As households decide optimally from a life-cycle perspective on consumption and the allocation of spouses' time to work, leisure, and child care, the model features all three ingredients our inequality measure seeks to incorporate (dynamics, multiple dimensions, and stochastic outcomes). In particular, state variables include household wealth and individual stochastic male and female wages. These are considered "transferable" (or redistributable), whereas demographic states such as age and the presence of young children are not. As a separate contribution, we show how Chebyshev polynomials can be used to approximate the value function in order to substantially lessen the computational burden. Using the estimated value functions for the household, our two-step construction determines first inequality for each transferable state, and then aggregates these univariate inequality measures using the shadow values as social weights.

The empirical application not only enables us to illustrate the construction of the new inequality index in a specific setting but also to make explicit and then to quantify important differences between dynamic and static inequality measurement. Unlike the static perspective, in which transferable states alone determine the value of the inequality measure, the dynamic perspective also incorporates the actions – and hence the resulting future states – that emerge. In particular, a static view on the Blundell et al. (2018) data shows dramatic differences between the inequality levels of wealth and husbands' wages. Adopting a dynamic perspective, the inequality gap is much reduced and husbands' wages are attributed much larger social weights than assets, wages being the key determinant of future consumption capacity and thus long-term well-being. This structural approach also enables us to evaluate the inequality impacts of policy interventions, since the behavioral responses of individuals are taken into account. For instance, we quantify income effects when wealth is changed across the wealth distribution, and examine the distributional consequences of a universal child care policy (via removing a fixed utility cost of work), leading to a reduction in female earnings inequality as female labor supply increases.

In the third part of the paper, we demonstrate that our inequality measure can be obtained using a reduced-form approach. Hence, a detailed structural modeling can be avoided, which widens the economic settings that can be analyzed. This follows since the inequality measure only makes use of the individual value functions, and not of the policy functions or the law of motion for states. Estimation of the value functions is

achieved by the deployment of offline reinforcement learning techniques, specifically offline Monte Carlo prediction. Given individual per-period utility function (also assumed in the classic static univariate Atkinson measure) and panel data that contain full trajectories of agents' states and actions, we can learn policy-dependent value functions. If the sampled trajectories are produced by optimally behaving agents, then the policy-dependent value function estimate resembles closely the optimal (i.e. "true") value function. In order to validate this learning approach in our empirical setting, we use the structural household life-cycle model with child care as our test laboratory. In practice, the approximation of the value function is obtained via separate feedforward, fully connected neural networks for households with and without children. We show that in this empirical setting reduced-form and structural approaches yield almost identical inequality measures.

The paper is structured in three parts. First, in Section 2 we present the inequality measurement approach in general form. The construction is then illustrated in Section 3 in the context of a specific setting, the household life-cycle model with children, and applied to US data for married couples taken from the PSID. This setting is also used to identify and quantity key differences between a pure static and univariate measurement of inequality, and our dynamic and multivariate approach. The reduced-form approach based on offline reinforcement learning in Section 4 constitutes the third part of the paper. For the sake of concreteness, the discussion is developed in the context of the empirical setting in the second part, and the juxtaposition of reduced-form and structural results enable us to validate our approach. Lastly, Section 5 concludes.

2. Inequality measurement approach

We adopt the terminology of dynamic optimization and distinguish between state and action variables. The state variables (or "states") fully characterize the current economic situation of an individual. They are further differentiated into transferable and non-transferable states. A variable is called transferable if it is, in principle, possible to shift some amount between individuals. Typical transferable states are wealth or income, since they could hypothetically be redistributed. Examples of non-transferable state variables are age or household composition.

2.1. The dynamic social evaluation function

We introduce the following notation. Let S be the number of transferable states, H the number of non-transferable states, N the number of individuals, and T the number of (discrete) time periods. The time horizon may be finite or infinite. We denote the value of the transferable state $k = 1, \ldots, S$ describing the situation of individual $i = 1, \ldots, N$ in period $t = 1, \ldots, T$ by $s^k_{t,i}$. We define the vector of states in period t of individual i as

$$\mathbf{s}_{t,i} = (s^1_{t,i}, \ldots, s^S_{t,i})$$

and the vector of state variable k in period t of all individuals as

$$\mathbf{s}^k_t = (s^k_{t,1}, \ldots, s^k_{t,N}).$$

In the same way, the denote the non-transferable states by $h^l_{t,i}$ for $l = 1, \ldots, H$ and define

$$\begin{aligned} \mathbf{h}_{t,i} &= (h^1_{t,i}, \ldots, h^H_{t,i}) \\ \mathbf{h}^l_t &= (h^l_{t,1}, \ldots, h^l_{t,N}). \end{aligned}$$

While the states of individual i in the current period $t = 1$ are given and known, their future values may depend (deterministically or stochastically) on actions that the individual takes in response to the current period's states. Denote the action variables by $a^m_{t,i}$ for $m = 1, \ldots, A$, where A is the number of action variables. The vector of action variables is $\mathbf{a}_{t,i}$. Common action variables are the amount of consumption, savings, time allocation, or portfolio allocation.

Taking the individual's actions into account is essential, as we assume that each individual can contribute to its own fortune by optimally choosing their actions, thus driving the evolution of their state variables. Formally, the states evolve according to the law of motion

$$\begin{bmatrix} \mathbf{s}_{t+1,i} \\ \mathbf{h}_{t+1,i} \end{bmatrix} = m(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}, \mathbf{a}_{t,i}, \boldsymbol{\varepsilon}_{t+1,i}) \tag{1}$$

where $\boldsymbol{\varepsilon}_{t+1,i}$ is a random shock vector. For simplicity, we assume that the model dynamics are solely ruled by the individual's own actions, the current states, and the random shocks.

Consequently, state or action variables of other individuals do not enter the law of motion.[2]

Each individual maximizes their expected discounted life-time return

$$U_i = E_1\left(\sum_{t=1}^{T} \beta^{t-1} u(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}, \mathbf{a}_{t,i})\right) \tag{2}$$

where E_1 is the conditional expectation operator given the information at time $t = 1$ and β is a discount factor. The function $u(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}, \mathbf{a}_{t,i})$ is a reward (or utility) function. The value function is defined as the maximum of U_i over the action space[3]

$$V(\mathbf{s}_{1,i}, \mathbf{h}_{1,i}) = \max_{\{\mathbf{a}_{t,i}\}_{t=1}^{T}} \left[E_1\left(\sum_{t=1}^{T} \beta^{t-1} u(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}, \mathbf{a}_{t,i})\right)\right] \tag{3}$$

subject to the law of motion (1). This can be reformulated recursively (Bellman 1957, 1962)

$$V(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}) = \max_{\mathbf{a}_{t,i}} [u(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}, \mathbf{a}_{t,i}) + \beta E_1(V(\mathbf{s}_{t+1,i}, \mathbf{h}_{t+1,i}))], \tag{4}$$

where $\mathbf{s}_{t+1,i}$ and $\mathbf{h}_{t+1,i}$ are determined according to the law of motion (1). The functional form of the value function depends on the specification of the law of motion, the instantaneous reward (or utility) u and the discount factor β. Negative state values are possible if the value function is set up accordingly. Uncertainty is included through the random shocks in the law of motion.

The individual value functions are aggregated into the dynamic social evaluation function SEF

$$SEF(\mathbf{s}_1^1, \ldots, \mathbf{s}_1^K, \mathbf{h}_1^1, \ldots, \mathbf{h}_1^H) = \frac{1}{N} \sum_{i=1}^{N} V(\mathbf{s}_{1,i}, \mathbf{h}_{1,i}). \tag{5}$$

We make a few technical assumptions about the SEF. It is assumed that SEF is continuous and differentiable in each transferable state variable. Non-transferable states are often categorial variables, representing, for example, the household composition or the health status of an individual. The SEF is anonymous in the sense that permuting

[2] In multivariate settings some states might represent macroeconomic variables, but in that case one may assume that the effect of individual actions on macroeconomic variables is infinitesimal small and can be ignored (at least from the point of view of an individual). In that sense, the macroeconomic environment is regarded as exogenous.

[3] For finite time horizons the value function typically depends on the period under consideration. To simplify notation, we do not explicitly indicate this dependence.

the individuals does not change the value of SEF. We further assume that SEF is monotone and (quasi-)concave in each transferable state variable. This holds for fairly broad conditions on the equation of motion (1) and the reward function (Benveniste and Scheinkman 1979).

2.2. The inequality index

In a first step, we consider inequality separately in each transferable state variable. As in Atkinson (1970), inequality in state variable $\mathbf{s}_1^k$ is measured by comparing total resources (i.e. the sum $\sum_i s_{1,i}^k$) with the minimum amount of total resources required to achieve the observed level of welfare, i.e. $SEF(\mathbf{s}_1^1, \ldots, \mathbf{s}_1^K, \mathbf{h}_1^1, \ldots, \mathbf{h}_1^H)$. That is, unless the marginal effects of the k^{th} transferable state on the SEF are all equal for the actual state distribution, one could attain the same level of social welfare with a smaller total amount of the k^{th} transferable state variable. Note that in the multidimensional setting the marginal effects generally depend on all state variables, not just on the one that is transferred. For example, the marginal effect of wealth (at a given level of wealth) is likely smaller in good health than in poor health. If the wealth level of the healthy person is decreased by one unit, the less healthy person needs less than one additional unit to keep the total level of social welfare constant. As a result, it might happen that wealth (as the transferable variable) is shifted away from a poor individual, if that person enjoys high levels in other state dimensions (e.g. good health).

Solving for the minimal required amount of total resources implies that individuals with the same non-transferable states $\mathbf{h}_{1,i}$ and the same non-transferred transferable states $\mathbf{s}_{1,i}^{-k} = (\ldots, s_{t,i}^{k-1}, s_{t,i}^{k+1}, \ldots)$ have the same level of the transferable state $\tilde{s}_{1,i}^k$. However, there can be substantial differences in the transferable state variable $\tilde{s}_{1,i}^k$ across individuals with diverse features $\mathbf{h}_{t,i}$ (e.g. different health or employment status) and $\mathbf{s}_{1,i}^{-k}$ (e.g. different amounts of wealth or income).

Formalizing the general idea laid out above, we solve the following nonlinear minimization problem,

$$\min_{\{\tilde{s}_{1,i}^k\}_{i=1}^N} \sum_{i=1}^N \tilde{s}_{1,i}^k \qquad \text{s.t.} \quad \widetilde{SEF}^k = SEF, \tag{6}$$

with

$$\widetilde{SEF}^k = SEF(\mathbf{s}_1^1, \ldots, \tilde{\mathbf{s}}_1^k, \ldots, \mathbf{s}_1^S, \mathbf{h}_1^1, \ldots, \mathbf{h}_1^H)$$
$$SEF = SEF(\mathbf{s}_1^1, \ldots, \mathbf{s}_1^k, \ldots, \mathbf{s}_1^S, \mathbf{h}_1^1, \ldots, \mathbf{h}_1^H).$$

Hence, only the distribution in the k^{th} state variable differs from the actually observed distribution. Additional inequality constraints can be imposed, if necessary, e.g. $\tilde{s}_{1,i}^k \geq \underline{s}_k$ for $i = 1, \ldots, N$ with $\underline{s}_k$ being a lower bound in dimension k. Except for variables for which the optional inequality constraints bind, all marginal effects on SEF must be equal at the optimum due to the assumed monotonically increasing and strictly quasi-concave shape of SEF. Hence, numerically solving the optimization problem is straightforward – either by transferring the k^{th} state variable from individuals with low marginal effects on SEF to individuals with high marginal effects in an iterative procedure until convergence or by directly searching for the optimal marginal effect on SEF through a bisection method and inferring the optimal state distribution from it.

To ensure that the equality constraint in (6) holds, some share of the transferred amount is "saved". Inequality in state dimension k is measured by the share of resources that could be saved if the state variable was distributed as $\tilde{\mathbf{s}}_1^k$ rather than $\mathbf{s}_1^k$.

As results of the S minimizations we obtain the S univariate inequality measures

$$I_k = 1 - \frac{\sum_{i=1}^N \tilde{s}_{1,i}^k}{\sum_{i=1}^N s_{1,i}^k}. \tag{7}$$

These indices are between 0 (no inequality) and 1 (highest inequality) if the numerator is non-negative, i.e. $\sum_{i=1}^N \tilde{s}_{1,i}^k \geq 0$, and the denominator is strictly positive, $\sum_{i=1}^N s_{1,i}^k > 0$. We assume that these conditions are always satisfied.

In a second step, we aggregate the S univariate inequality measures into a measure of total inequality. To do so, we assess the relative importance of each state by employing the "shadow values" that are associated with the constraints in (6). To derive the shadow

values we set up the S Lagrangians

$$\begin{aligned}\mathcal{L}_k(\tilde{s}^k_{1,i},\ldots,\tilde{s}^k_{1,N},\lambda^k,\mu^k_1,\ldots,\mu^k_N) = &\sum_{i=1}^{N}\tilde{s}^k_{1,i}\\ &-\lambda^k(\widetilde{SEF}^k - SEF)\\ &-\sum_{i=1}^{N}\mu^k_i(\tilde{s}^k_{i1}-\underline{s}_k),\end{aligned} \tag{8}$$

where λ^k is the shadow value for state variable $k = 1,\ldots,S$.[4] The first order conditions for the k^{th} state variable are

$$\frac{\partial\mathcal{L}_k}{\partial\tilde{s}^k_{1,i}} = 1-\lambda^k\frac{\partial\widetilde{SEF}^k}{\partial\tilde{s}^k_{1,i}}-\mu^k_i = 0, \tag{9}$$

for $i = 1,\ldots,N$. If the inequality constraint is not binding for individual i, then $\mu^k_i = 0$ and

$$\lambda^k = \left(\frac{\partial\widetilde{SEF}^k}{\partial\tilde{s}^k_{1,i}}\right)^{-1}. \tag{10}$$

The shadow value λ^k expresses the required change in the state $s^k_{1,i}$ per infinitesimal unit change in the constraint, i.e. if the minimum required welfare was relaxed or tightened by an infinitesimal welfare unit.[5] Hence, the shadow values help make the state dimensions – and their inequalities – comparable by introducing social welfare as the common metric. As a consequence, if the k^{th} state variable is of great importance for social welfare, its shadow value is smaller than that for the less important variables.

Once the shadow values $\lambda^1,\ldots,\lambda^S$ have been calculated according to (10), we aggregate the univariate inequality measures,

$$I = 1-\frac{\sum_{k=1}^{S}\frac{1}{\lambda^k}\sum_{i=1}^{N}\tilde{s}^k_{1,i}}{\sum_{k=1}^{S}\frac{1}{\lambda^k}\sum_{i=1}^{N}s^k_{1,i}} \tag{11}$$

The index gives the weighted maximum share of the total amount of state variables

[4]The shadow values of the inequality constraints, if any, are contained in the N-dimensional vector $\mu^k = (\mu^k_i,\ldots,\mu^k_N)$. If no inequality constraints are imposed, this part of the Lagrangian vanishes.

[5]The equality constraint is always binding due to the concavity assumption. Thus, it always holds that $\lambda_k \neq 0$ (see Karush (1939); Kuhn and Tucker (1951)).

that could have been saved by optimizing the distribution of states so that the current level of social welfare is still reached. Rewriting (11) yields

$$I = \sum_{k=1}^{S} \omega_k I_k \tag{12}$$

with

$$\omega_k = \frac{\frac{\sum_{i=1}^{N} s_{1,i}^k}{\lambda^k}}{\sum_{l=1}^{S} \frac{\sum_{i=1}^{N} s_{1,i}^l}{\lambda^l}}. \tag{13}$$

Thus, total inequality is a convex combination of the inequality in each state variable. The weights $\omega_1, \ldots, \omega_S$ are the shares of the state variables measured in social welfare units as the common metric. Equation (12) is a natural decomposition of total inequality into the contributions of all transferable state variables.

While the approach closely follows the ideas of Atkinson (1970) our inequality measure (11) takes into account economic mobility and uncertainty. In this respect, the dynamic inequality measure takes account of future state trajectories that might be inherently uncertain instead of considering a snapshot of the state distribution and deriving a static measure of inequality of outcomes.

The static environment is nested in our dynamic structure. If $S = T = 1$ and $H = 0$, it is easily shown that our inequality measure collapses to the so-called Atkinson index.[6]

3. Application: Inequality in a household life-cycle model with child care

Our approach to measuring inequality is structural, the basic ingredient being the individual value functions of a dynamic programming problem. To illustrate our approach in

[6]The SEF becomes $SEF(s_1, \ldots, s_N) = \frac{1}{N} \sum_{i=1}^{N} u(s_i)$ where s_i is the value of the single state variable (e.g. income) and u denotes the utility function. The minimization problem (6) can be solved analytically. The equally distributed equivalent (e.d.e.) $\tilde{s}$ is defined by the equality $SEF(\tilde{s}, \ldots, \tilde{s}) = SEF(s_1, \ldots, s_N)$, which implies that $\tilde{s} = u^{-1}\left(\frac{1}{N} \sum_{i=1}^{N} u(s_i)\right)$. Then, the resulting inequality measure is the Atkinson index given by

$$I(s) = 1 - \frac{\frac{1}{\lambda} \sum_i \tilde{s}}{\frac{1}{\lambda} \sum_i s_i} = 1 - \frac{\tilde{s}}{\frac{1}{N} \sum_i s_i}.$$

a specific setting, we consider the household life-cycle model with child care presented in Blundell et al. (2018). In the dynamic setting of the life cycle, they study households' choices on consumption and allocation of spouses' time to work, leisure, and child care. In the face of uncertainty, the allocation of assets and time over the life cycle serves the purpose of smoothing the marginal utility in response to shocks. Hence, the model features all the aspects that our inequality measure seeks to capture: multidimensionality through the state variables wages and assets, dynamics, and uncertainty via earnings risk. In our taxonomy of states, the model has three transferable state variables (i.e. assets and wages of both spouses) and two non-transferable states (i.e. age and a binary variable indicating whether young children live in the household).

3.1. The model

We briefly sketch a slightly simplified version of the model presented in Blundell et al. (2018). A household i optimizes its current expected discounted life-time utility (their "return" $U_{t,i}$ at age t) over the choice of consumption ($C_{t,i}$) and each spouses' allocation of time to labor ($H_{1,t,i}, H_{2,t,i}$), leisure ($L_{1,t,i}, L_{2,t,i}$), and household production ($T_{1,t,i}, T_{2,t,i}$) for all ages $t = 1, \cdots, T$. These actions are collected in the vectors $\mathbf{a}_{t,i} = (C_{t,i}, H_{1,t,i}, H_{2,t,i}, L_{1,t,i}, L_{2,t,i}, T_{1,t,i}, T_{2,t,i})$. The corresponding value function is

$$V(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}) = \max_{\{\mathbf{a}_{\tau,i}\}_{\tau=t}^{T}} \left[E_t \left(\sum_{\tau=t}^{T} \beta^{\tau-t} u(\mathbf{a}_{\tau,i}, \mathbf{h}_{\tau,i}) \right) \right], \tag{14}$$

where $\mathbf{s}_{t,i} = (A_{t,i}, W_{1,t,i}, W_{2,t,i})$ is the vector of transferable states consisting of the household's assets $A_{t,i}$ as well as the husband's and wife's wage rates ($W_{1,t,i}, W_{2,t,i}$) at age t. The vector of non-transferable states, made up of a binary young-children indicator ($z_{t,i} = 0$ if there are no children, $z_{t,i} = 1$ if there are children) and age t, is denoted by $\mathbf{h}_{t,i} = (z_{t,i}, t)$.[7]

[7]The household-specific subscript for age is dropped for notational clarity.

The household's contemporaneous utility function is given by

$$
\begin{aligned}
u(\mathbf{a}_{\tau,i}, \mathbf{h}_{\tau,i}) = \exp(\tilde{\phi}_C(z_{t,i}))\frac{\tilde{C}_{t,i}^{1-1/\eta}}{1-1/\eta} \\
- \frac{1}{1-\rho_L}\left[\exp(\tilde{\phi}_{L_1}(z_{t,i}))L_{1,t,i}^{1-1/\varphi_{L_1}} + \exp(\tilde{\phi}_{L_2}(z_{t,i}))L_{2,t,i}^{1-1/\varphi_{L_2}}\right]^{1-\rho_L} \\
- \frac{1}{1-\rho_T}\left[\exp(\tilde{\phi}_{T_1}(z_{t,i}))T_{1,t,i}^{1-1/\varphi_{T_1}} + \exp(\tilde{\phi}_{T_2}(z_{t,i}))T_{2,t,i}^{1-1/\varphi_{T_2}}\right]^{1-\rho_T},
\end{aligned} \tag{15}
$$

where $\tilde{C}_{t,i} = C_{t,i} - \gamma(z_{t,i})\mathbf{1}\{H_{2,t,i} > 0\}$ is accounting for the non-separability between household consumption and female employment through the fixed utility cost of work parameter $\gamma(z_{t,i})$ which differs for couples with and without children. Furthermore,

$$
\tilde{\phi}_x(z_{t,i}) = \phi_x^{nk} + \phi_x^k z_{t,i}
$$

for $x = \{C, L_1, L_2, T_1, T_2\}$. To ensure that the utility function is well-behaved, the following restrictions are imposed,

$$
0 < \varphi_x < 1 \text{ for } x = \{L_1, L_2, T_1, T_2\},\ \eta > 0, \text{ and } \rho_x < 1 \text{ for } x = \{L, T\}.
$$

The law of motion for assets $A_{t,i}$ is

$$
A_{t+1,i} = (1+r)\left[A_{t,i} + \mathcal{T}\left(z_{t,i}, H_{1,t,i}, H_{2,t,i}, W_{1,t,i}, W_{2,t,i}\right) - C_{t,i}\right],
$$

where r is the interest rate on risk-free assets. The household's net earnings are

$$
\mathcal{T}\left(z_{t,i}, H_{1,t,i}, H_{2,t,i}, W_{1,t,i}, W_{2,t,i}\right) \approx \mathcal{X}\left(b(z_{t,i}) + H_{1,t,i}W_{1,t,i} + H_{2,t,i}W_{2,t,i}\right)^{1-\mu}.
$$

The parameters $\mathcal{X}$ and μ characterize the tax function, and $b(z_{t,i})$ is an income floor, e.g. government benefit payments, that may depend on the household composition. Liquidity and time-use are constrained by

$$
\begin{aligned}
&A_{t,i} \geq 0, H_{j,t,i} \geq 0, L_{j,t,i} \geq 0, T_{j,t,i} \geq 0 \quad \forall t \\
&H_{j,t,i} + L_{j,t,i} + T_{j,t,i} = \bar{L} \text{ with } j = 1, 2,
\end{aligned}
$$

where $\bar{L}$ is the maximum time available to each spouse.

Uncertainty is introduced through the wage process for each spouse. The evolution of logarithmized wages is governed by

$$\log W_{j,t,i} = \mathbf{x}_t' \boldsymbol{\beta}_j^W + \log W_{j,t-1,i} + v_{j,t,i} \text{ with } j = 1, 2.$$

This wage process is subject to a permanent shock $v_{j,t,i}$. Mean logarithmized wages are a function of observable characteristics $\mathbf{x}_{j,t,i}$ weighted by the corresponding parameter vector $\boldsymbol{\beta}_j^W$. While the shocks are uncorrelated between households and over time, they are correlated between spouses. In a first estimation step, Blundell et al. (2018) find the spouse-specific mean age-profiles of wages over the life-cycle by a quadratic regression, i.e. $\mathbf{x}_t = (1, t, t^2)'$.

3.2. Computing the value function and the inequality index components

Our inequality measurement approach requires solving the value function (14) given the reported structural parameters.[8] First, we observe that the computational burden can be significantly reduced by approximating the value function with Chebyshev polynomials, as proposed in Cai and Judd (2013). These efficiency gains arise because the value function is continuous and concave.[9] Moreover, the resulting value function is differentiable over the entire domain, a requirement for our measurement approach.

The remaining details of the numerical implementation are as follows. The asset space and the logarithmized wage spaces are discretized with 50 and 15 points, respectively. We assume that individuals enter the labor market at the age of 25 and leave it at 65 years. As in Blundell et al. (2018) the presence of young children deterministically depends on age, that is half of the women give birth at the age of 28. Thus, we solve the value functions for both spouses who have young children and spouses who do not. Children

[8] See Blundell et al. (2018), Table 2, p. S93, and Blundell et al. (2019), Table 3A p. 2561, Table 3B, p. 2561.

[9] The standard numerical method is piecewise linear interpolation, which requires many grid nodes because the value function is continuous and concave. Otherwise, the approximation is likely to be far off the true value function, which in turn leads to suboptimal results.

are assumed to have an impact on taste shifters and utility of consumption as long as they are young, that is, for couples with children $z_{t,i} = \mathbf{1}\{28 \leq t \leq 37\}$.

In order to avoid optimizing the value function over seven action dimensions in each backward induction iteration, we solve for conditionally optimal leisure and household production levels given the values of both labor variables before running the dynamic programming algorithm.

Then, given the value functions and the data of the observed states, we can determine the optimized state distribution and the respective Lagrange multiplier in each state dimension. The univariate dynamic inequality measures and weights are stated in equations (7) and (13). We proceed by searching for the partial derivative of the SEF in the k^{th} state dimension that corresponds to a state distribution conforming with the constraint on the SEF in equation (6). In practice, we solve this problem through a bisection method that utilizes unidimensional cuts through the SEF in the k^{th} dimension and an inverse gradient function, which returns a state distribution for the SEF's gradient. The whole procedure is outlined in Algorithm 1 in Appendix A. The logic behind this method is that all partial derivatives must be equal in an optimized state distribution or else one could further decrease the total amount of the state variable by shifting some amount of the state from an individual with a low to one with a high partial effect on the SEF. An exception are individuals, for whom the state variable is constrained by the lower state bound. Hence, to resolve this issue we extrapolate the inverse gradient function $f(\cdot)$ by setting $f(x) = \underline{s}_k$ for all $x \leq \partial sef(\mathbf{s}_1^k)/\partial \underline{s}_k$, where $sef(\mathbf{s}^k)$ is a unidimensional cut through the social evaluation function $sef(\mathbf{s}^k) = SEF(\mathbf{s}^k|\mathbf{s}_1^1, \ldots, \mathbf{s}_1^{k-1}, \mathbf{s}_1^{k+1} \ldots, \mathbf{s}_1^K, \mathbf{h}_1^1, \ldots, \mathbf{h}_1^H)$. To ensure that the bisection method finds a global optimum, we must impose theoretically well-founded monotonicity and concavity constraints on the unidimensional cuts through the value function. Hence, we determine those by shape-preserving Chebyshev approximation (Cai and Judd 2013).

3.3. Data: Couples in the USA

Our data of married couples in the US are obtained from the *Panel Survey of Income Dynamics* (PSID). We follow Blundell et al. (2018), and consider a sample of all married couples in which the wife is between 25 and 65 years old and in which no spouse earns a

wage below half of the minimum wage. For simplicity, we "anchor" the inequality analysis in the year 2000, in order to allow for sufficiently long time horizons, yielding a sample size of 2,199 couples. The year 2000 is also the first survey year that contains detailed asset holding and consumption expenditure data in addition to the usual income and demographic data. (Average) Hourly wages are obtained by dividing the reported annual individual labor earnings by the reported annual hours of work.

Table 1 provides some descriptive statistics. As expected, men work longer hours and have higher hourly wages than women. Husbands' and wives' median annual hours of work are similar to their mean, with the differences in means across the sexes being 532 hours. The distribution of wages (in 2000 prices) is skewed. The distribution of assets across couples is highly skewed, the mean differing from the median by a factor of 2.8. Table 2 provides the correlation matrix, and shows the expected positive correlations between assets and wages, and the wages of each spouse. Correlations between spousal wages exceed those between assets each spousal wage.

Table 1: Married couples in the US in 2000: Descriptive statistics

Statistic	Mean	Pctl(25)	Median	Pctl(75)	Atkinson
Hours of husband	2265	1980	2160	2548	
Hours of wife	1733	1396	1900	2056	
Wage of husband	22.90	12.73	18.45	27.29	0.10
Wage of wife	16.34	8.93	13.17	20.26	0.09
Assets	235387	30000	85000	215000	0.38

Notes. The sample contains all employed married couples in the 2000 wave of the PSID dataset, where the wife is between 25-65 years old. Wages are in 2000 prices. "Atkinson" refers to the (static univariate) Atkinson inequality index (see Footnote 6 for definitions). Number of observations/couples: $n = 2,199$.

Our dynamic inequality analysis focuses on wealth and the earnings of husbands and wives. As a benchmark for our dynamic measure, we report in the last column of Table 1 the *static and marginal* Atkinson inequality indices (see Footnote 6), based on the commonly deployed utility function $u(s^k_{t,i}) = ({s^k_{t,i}}^{1-\gamma} - 1)/(1-\gamma)$ and the inequality aversion parameter $\gamma = 0.5$. We observe high inequality in the asset dimension. In contrast, wage inequality is much lower and does not differ much between male and female wages.

Table 2: Correlation Matrix

	Assets	Wage of husband	Wage of wife
Assets	1	0.286	0.213
Wage of husband	0.286	1	0.360
Wage of wife	0.213	0.360	1

Finally, we illustrate how the data are taken to the model by depicting section cuts of the couple's value function given by Equation (14). To this end, we consider for Figure 1 a household with children and a wife aged 30. The left panel shows the cut through the value function where the female wage is set to the observed median of female wages, while the right panel displays the cut where the asset level equals the observed median of assets. The value functions are well-behaved, concave, and show a steep increase in a couple's maximized expected life-time utility when assets are low and male wages rise from low levels. By contrast, the increase in utility is less steep for low male wages as assets increase from low levels.

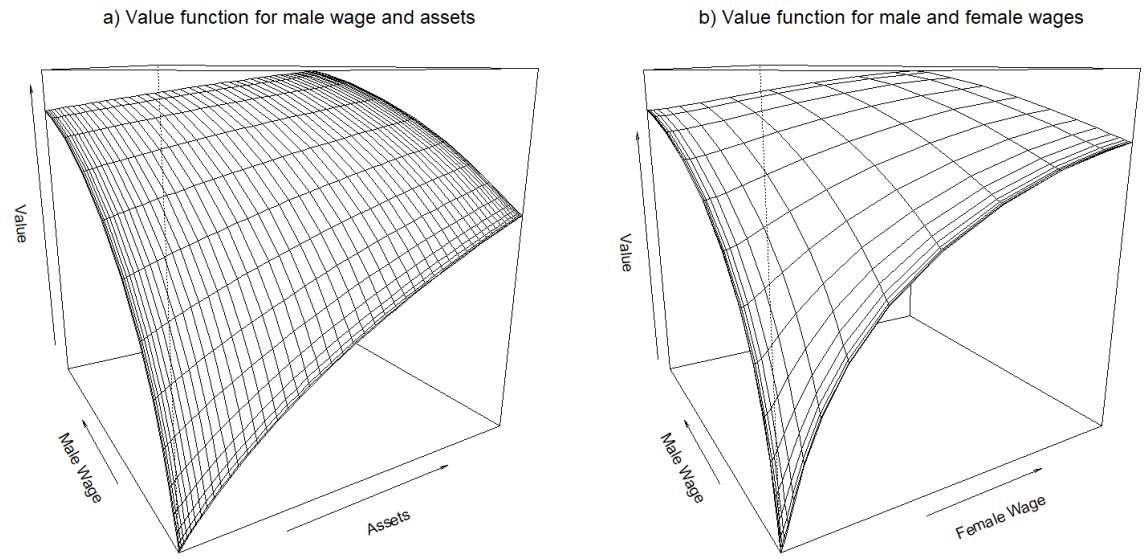
3.4. Measuring inequality dynamically

Given the value functions and the data of the observed states, we determine the optimized state distribution and the respective Lagrange multiplier in each state dimension, as detailed in Section 3.2. We explicitly consider assets and wages of husbands and wives, i.e. $\mathbf{s}_{t,i} = (A_{t,i}, W_{1,t,i}, W_{2,t,i})$. Table 3 reports the components of the dynamic inequality measure. The univariate measures for each state are then aggregated according to equation (11) to produce our dynamic inequality index:

$$I = I_A \cdot \omega_A + I_{W_1} \cdot \omega_{W_1} + I_{W_2} \cdot \omega_{W_2} \approx 0.4329$$

We proceed to discuss the results in Table 3, specifically in comparison to the static inequality measures of Table 1. Unlike in static inequality measurement, the states in our dynamic setting do not exclusively determine the inequality measure's value but also the future actions which in turn determine the future states, leading to important differences between dynamic and static inequality measures.

Figure 1: Value functions



Notes. The left panel shows the three-dimensional cut through the value function (VF) when the female wage is set to the median of observed female wages, the right panel depicts the cut through the VF when the asset level is set to the observed median asset level. The non-transferable state variables are set to $z = 1$ (i.e. at least one young child) and wife's age $t = 30$.

The asset dimension exhibits again the highest inequality. However, the difference between asset and wage dimensions is no longer as stark as before, since we take account of lifetime state trajectories, and wages are an important determinant of future consumption capacity. Therefore, contemporaneous differences in wages have a high impact on the expected lifetime return, and inequality in wages should be given a higher weight, even if the static wage inequality is low. In particular, the husband's wage is given the highest weight, which is consistent with the persistent gender roles in the United States with the husband as breadwinner (see also Table 1). The univariate inequality index of the wife's wages also has a higher effect on composite inequality than assets, indicating that a households' current position in the bivariate wage distribution is rather persistent and indicative of the lifetime return due to the compounding effect of income processes.

Next, we highlight another interesting feature captured by our new methodology. Although the static marginal Atkinson indices in Table 1 assess the wage inequality to

Table 3: Estimation Results: Inequality Components

	A	W_1	W_2
$I_{\{\cdot\}}$	0.55	0.36	0.48
$\omega_{\{\cdot\}}$	0.17	0.50	0.33

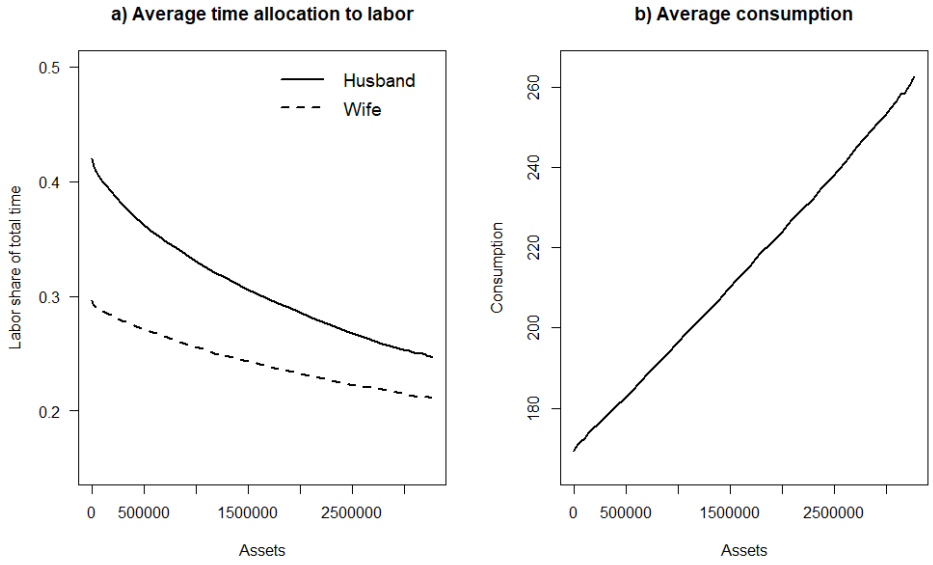
Notes. The univariate dynamic inequality measures ($I_{\{\cdot\}}$) and weights ($\omega_{\{\cdot\}}$) are defined in equations (7) and (13) for assets (A) and wages for husband (W_1) and wife (W_2) in the married couple. The univariate measures are then aggregated according to equation (11) to yield $I = I_A \cdot \omega_A + I_{W_1} \cdot \omega_{W_1} + I_{W_2} \cdot \omega_{W_2} \approx 0.4329$.

be approximately equal for husbands and wives, we find that dynamic wage inequality is significantly higher for wives than for husbands. The reason for this difference is the fixed utility-cost of work component $\gamma(z_{t,i})$, which reduces household utility if the wife is in employment. Hence, women with low wages have an incentive not to enter the labor market, and this in turn reinforces the dynamic inequality between households with high and low female wages. We verify this reasoning in Section 3.5 below, where we remove the fixed cost of female employment in a counterfactual policy simulation experiment.

In order to examine the dependence of the inequality measure across dimensions, we let the asset distribution change exogenously by multiplying all assets by a scaling factor $\xi = \{0.1, 0.5, 1\}$. Note that most static inequality measures (such as the Gini coefficient) do not react to such a change as they are scale invariant. Figure 2 visualizes how asset level and labor supply decisions are intertwined. Panel a) shows the decrease in the labor supply of both husband and wife, as the couple becomes richer. At the other extreme, as a household becomes asset poor, the husband increases his labor supply more steeply than the wife. Panel b) depicts the increase in consumption as assets increase.

Table 4a reports the results for measured inequality. When the asset distribution is scaled up, the inequality in the asset dimension *decreases*. This seemingly counter-intuitive result arises because of dynamic linkages. Since assets and wages are positively correlated (see Table 2), asset-poor households tend to earn low wages. The value function evaluated at low assets and wages has thus a steep slope. Therefore, reducing the state in one dimension results in a large decrease in value. On the other hand, asset-rich households are likely to also have high wages. The value function in these areas is already relatively flat, thus a reduction of one state does not cause a large decrease in value for such households. Although rich households lose more in absolute terms, the effect with regard to lifetime

Figure 2: Policy reactions: Changes in assets.



Notes. The optimal policy is averaged over both wage dimensions, family composition, and age. The figure quantifies the effects induced by changing assets exogenously.

utility is much more pronounced for poor households. Consequently, more assets could be transferred while keeping social welfare constant. Hence, inequality rises in the this state dimension. However, we also point out that the weight of the asset dimension decreases with a reduced scale factor, because the lower level of assets makes them less relevant for overall inequality.

In contrast, the weights in both wage dimensions rise, and the wages' univariate inequalities decrease with a reduction in the dispersion of the asset distribution. Reversing the reasoning from above, this result is quite intuitive. Asset-poor households now react even stronger to increases in the wage dimensions as they are located in the steeper area of the value function, while the rich households are less ignorant with respect to reductions in the wage dimensions as the reduction in asset has moved them closer to the steep area of the value function. Thus, less wages can be transferred between households and inequality in the wage dimensions decreases with a smaller dispersion in the asset dimension.

Table 4: Effects of Changes to the Asset Distribution

(a) Inequality measures

ξ	I	I_A	I_{W_1}	I_{W_2}
0.1	0.35	0.65	0.30	0.42
0.5	0.40	0.58	0.33	0.45
1	0.43	0.55	0.36	0.48

(b) Weights

ξ	ω_A	ω_{W_1}	ω_{W_2}
0.1	0.03	0.59	0.38
0.5	0.11	0.54	0.35
1	0.17	0.50	0.33

Notes. ξ is the scaling factor applied uniformly to assets. The univariate measures (I_A, I_{W_1}, I_{W_2}) are aggregated according to equation (11) to yield the dynamic measure I. The policy reactions are depicted in Figure 2.

3.5. Inequality comparisons for counterfactual policies

The structural approach to inequality measurement enables us to credibly evaluate the long-term inequality impact of policy interventions, since not only their direct redistributive effects, but also the behavioral responses of individuals are taken into account when determining the value function. Changes in the value function in turn have an impact on the solution of the minimization problem and, as a consequence, on measured inequality.

We proceed to evaluate a policy measure that sets the fixed utility-cost of work component $\gamma(z_{t,i})$ to zero for all households. This might be achieved by offering free universal child care or government benefits for double-income households. We solve the value functions of the problem with $\gamma(z_{t,i}) = 0$ for $z_{t,i} = \{0, 1\}$ and determine the inequality measures as outlined above. The resulting counterfactual univariate inequality measures are reported in Table 5. Aggregating the univariate inequality indices yields the total inequality index:

$$I = I_A \cdot \omega_A + I_{W_1} \cdot \omega_{W_1} + I_{W_2} \cdot \omega_{W_2} \approx 0.39.$$

Comparing the results to Table 3, we observe that the policy measure leads to a decrease in total inequality. In particular, while male wage inequality is unaffected, the wife's wage inequality is now sharply reduced, becoming similar to male wage inequality. This evidence strengthens our argument that the high inequality of female wages is mainly

Table 5: Counterfactual inequalities: No fixed utility-cost of work

	A	W_1	W_2
$I_{\{\cdot\}}$	0.52	0.36	0.38
$\omega_{\{\cdot\}}$	0.17	0.48	0.35

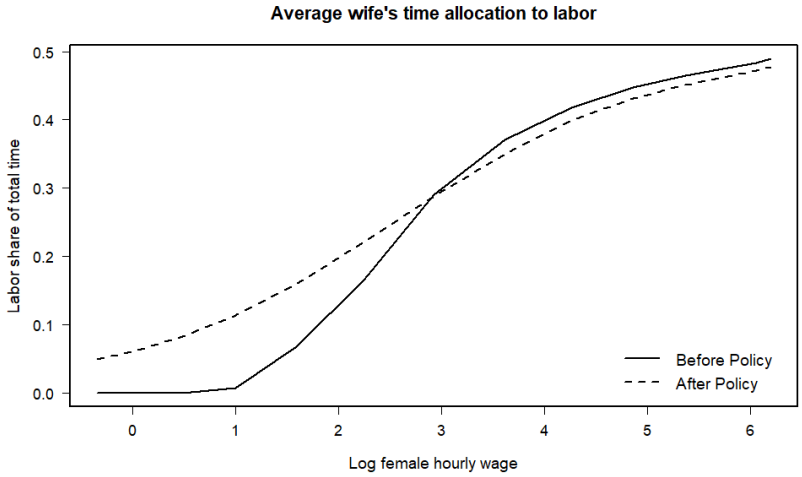
Notes. As per Table 3. The univariate measures are aggregated to yield $I = I_A \cdot \omega_A + I_{W_1} \cdot \omega_{W_1} + I_{W_2} \cdot \omega_{W_2} \approx 0.39$. The counterfactual policy imposes $\gamma(z_{t,i}) = 0$ for $z_{t,i} = \{0, 1\}$, i.e. no utility-cost of work irrespective of children in the household.

due to the fixed utility-cost of work component. This variable makes low-paying jobs unprofitable for women as the fixed cost of employment is incurred once such a job is taken up. Hence, in households where the wife's wage level is low, it is only the husband who is employed, resulting in a relatively low expected lifetime return for the couple. This incentive effect is also reflected in the policy functions of households. Figure 3 shows the wives' relative time allocation to labor for different levels of female wages (averaged over all ages, family compositions, asset levels, and male wage levels). The labor share is displayed both before and after the policy is implemented. With a fixed utility-cost of work, wives stay unemployed at low wage levels. After introducing the policy, they enter the labor market. Another effect of the policy is an increase in the weight of female wages ω_{W_2}. This is due to the women's increased relevance for household earnings once the female employment numbers rise. At the same time, measured inequality in assets as well as the weight ω_A also decrease slightly owing to the policy, which reflects the decreased long-term relevance of assets for consumption smoothing as the earnings of low-female-wage households increase after the policy's induction.

4. A reduced-form approach to measuring inequality based on Offline Reinforcement Learning

While Section 3.1 details the construction of the inequality measure in the specific setting of Blundell et al. (2018), it may initially appear to require structural econometric estimation. However, this is not the case, as the dynamic inequality measure can be computed using a reduced-form approach. This is possible because the measure only

Figure 3: Female work hours and the utility cost of work



Notes. The figure depicts the change in the wives' (log) hours allocated to paid work in response to policy of reducing fixed utility-cost of work component to zero (i.e. $\gamma(z_{t,i}) = 0$ for $z_{t,i} = \{0, 1\}$).

depends on individual value functions, but neither on the policy functions nor the law of motion. Consequently, our inequality measurement approach can be applied even in complex economic environments where the value function cannot be derived numerically (or analytically). The trade-off of the reduced-form approach is that it requires more extensive data. Apart from the contemporaneous utility function (also assumed to be known for the classic static univariate Atkinson measure), one has to observe panel data that contain trajectories of agents' states and actions. In the next section, we demonstrate how to learn the value functions if these data are available.

4.1. Offline Monte Carlo Prediction

Without an underlying structural model, traditional dynamic programming methods for computing the value function are not applicable. This is due to the lack of information regarding the evolution of states when the law of motion (1) is unknown. However, given panel data on actions and states, the value function can be estimated by offline reinforcement learning (Lange et al. 2012; Sutton and Barto 2018; Levine et al. 2020).

To this end, we generalize the value function (3), which gives the expected discounted lifetime return for the optimal choice of actions, to a policy-dependent value function as

defined in Sutton and Barto (2018, p. 59)

$$V_\pi(\mathbf{s_{t,i}}, \mathbf{h_{t,i}}) = E_\pi\left(U_{t,i}\right), \tag{16}$$

where $\pi : S \times H \to A$ is a policy, i.e. a mapping from states to actions. $U_{t,i}$ is the expected discounted return as in (2) starting at age t. The policy-dependent value function can be written as

$$\begin{aligned} V_\pi(\mathbf{s_{t,i}}, \mathbf{h_{t,i}}) &= E_\pi\left(\sum_{\tau=t}^{T} \beta^{\tau-t} u(\mathbf{s}_{\tau,i}, \mathbf{h}_{\tau,i}, \mathbf{a}_{\tau,i})\right) \\ &= u(\mathbf{s}_{t,i}, \mathbf{h}_{t,i}, \pi(\mathbf{a}_{t,i}|\mathbf{s}_{t,i}, \mathbf{h}_{t,i})) + \beta E_\pi(V_\pi(\mathbf{s}_{t+1,i}, \mathbf{h}_{t+1,i})). \end{aligned} \tag{17}$$

When the agent behaves optimally, denoted by policy π^*, equation (17) corresponds to equation (4).

Given a panel data set $\mathcal{D} = \{(\mathbf{s_{t,i}}, \mathbf{h_{t,i}}, \mathbf{a_{t,i}}, \mathbf{s_{t+1,i}}, \mathbf{h_{t+1,i}})\}$ that contains full trajectories of agents' states and actions for $t = 1, \ldots, T$, and a contemporaneous utility function $u(\mathbf{s_{t,i}}, \mathbf{h_{t,i}}, \mathbf{a_{t,i}})$, we can approximate the policy-dependent value function (17) by the conditional sample averages

$$V_\pi(\mathbf{s}, \mathbf{h}) \approx \hat{V}_\pi(\mathbf{s}, \mathbf{h}) = \hat{E}\left(U_{t,i}|\mathbf{s}_{t,i} = \mathbf{s}, \mathbf{h}_{t,i} = \mathbf{h}\right). \tag{18}$$

If the agents' behavior is optimal, and if the sampled trajectories cover a wide range of the state spaces, the solution of (18) approximates the optimal value function (4).[10] Using an appropriate function approximation method, we can estimate the policy-dependent value function for a wide range of states.

[10] This type of reinforcement learning is called offline Monte Carlo prediction as we use a fixed dataset of sampled trajectories to estimate the state value of any observed policy (Sutton and Barto 2018, p. 92). See e.g. Levine et al. (2020) for a discussion of the recent offline reinforcement learning literature concerned with finding optimal policies from static datasets of trajectories.

4.2. Validation in the Lab: Reduced-form inequality measurement in the household life-cycle model

In order to validate the reduced-form approach, we run a laboratory experiment, using the setting of Section 3.1. Since there are no real-world datasets providing panel data on both actions and states, we augment the PSID data from 2000 (providing the state variables) by simulating actions and future states for all couples up to (wife's) age 65. The resulting artificial panel dataset comprises 54013 observations on 2199 households. The inequality measured by offline reinforcement learning can then be compared (and should be similar) to the results of the structural approach.

The estimation procedure is detailed in Algorithm 2 in Appendix A.2; details regarding the deep learning architecture and its hyperparameter tuning routine are collected in Appendix A.3. Given the simulated trajectories of states and actions, we calculate the contemporaneous rewards in each period for each household according to equation (15). As the households' trajectories cover the entire remaining observation horizon, we can simply add the discounted rewards to obtain the lifetime utilities (returns $U_{t,i}$). If a panel dataset contains non-terminated trajectories, the lifetime return could be approximated by predicting the remaining cumulated discounted rewards by the policy-dependent value function estimate; see (17).

To approximate the value function and be able to generalize to unobserved states, we use separate feed-forward, fully connected neural networks for households with and without children. Following Runje and Shankaranarayana (2023), we square the weights of the neural network and choose a bounded exponential linear unit as activation function of the input and hidden layers in order to ensure that the value function increases monotonically.

Having estimated the value functions, we can proceed as before and construct our inequality measure. We obtain the univariate inequality measures and respective weights shown in Table 6. Total inequality is then

$$I = I_A \cdot \omega_A + I_{W_1} \cdot \omega_{W_1} + I_{W_2} \cdot \omega_{W_2} \approx 0.401$$

All results are close to the results in Table 3, indicating that a reduced form estimation of the inequality measure is a valid alternative to the structural approach.

Table 6: Reduced-Form Estimation Results

	A	W_1	W_2
$I_{\{\cdot\}}$	0.58	0.32	0.40
$\omega_{\{\cdot\}}$	0.21	0.46	0.33

Notes. The table shows that reduced-form and structural estimation yield essentially the same results. The univariate measures are aggregated to yield $I = I_A \cdot \omega_A + I_{W_1} \cdot \omega_{W_1} + I_{W_2} \cdot \omega_{W_2} \approx 0.401$. The results for the structural approach are reported in Table 3.

5. Conclusion

In this paper we have proposed a novel unifying approach to measure economic inequality in a dynamic, multidimensional setting with uncertainty, based on a dynamic social evaluation function. The inequality measure is based on the minimization of the total resources necessary to keep current social welfare constant. The shadow values of transferable state variables (e.g. earnings and wealth) enable us to weight the impact of each variable on economic inequality and to unify several variables in a single welfarist index of economic inequality. Our analysis in the context of a structural household life-cycle model has shown the merit of our measurement approach when compared to the static univariate approach prevelant in the inequality literature.

The structural model has been deployed for two objectives. First, using the structure of the model, we have estimated the value functions. These then allow the construction of the dynamic social evaluation function, and hence our inequality measure in a specific setting. The second objective is the set-up of a laboratory, in which we can validate our reduced-form approach based on offline reinforcement learning techniques, that requires only panel data for agents' states and actions, and the specification of the per-period utility function. Within this lab setting, we have verified that reduced-form and structurally estimated inequality are indeed very close. As a consequence, the reduced-form approach and thus our inequality measure can be deployed irrespective of the complexity of the

economic setting. Furthermore, the reduced-form approach could be used to perform statistical inference by standard bootstrap methods.

References

Aaberge, R. and A. Brandolini (2015). Multidimensional poverty and inequality. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of income distribution*, Volume 2 of *Handbooks in Economics*, Chapter 3, pp. 141–216. Amsterdam, Netherlands: North-Holland.

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

Bellman, R. (1957). *Dynamic Programming* (1 ed.). Princeton, NJ: Princeton University Press.

Bellman, R. E. (1962). *Applied Dynamic Programming*. Princeton Legacy Library. Princeton, NJ: Princeton University Press.

Benveniste, L. M. and J. A. Scheinkman (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica 47*(3), 727.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2018). Children, time allocation, and consumption insurance. *Journal of Political Economy 126*(S1), S73–S115.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2019). Erratum: Children, time allocation, and consumption insurance. *Journal of Political Economy 127*(5), 2559–2567.

Burbidge, J. B., L. Magee, and A. L. Robb (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association 83*(401), 123–127.

Cai, Y. and K. L. Judd (2013). Shape-preserving dynamic programming. *Mathematical Methods of Operations Research 77*(3), 407–421.

Cowell, F. A. (2000). Measurement of inequality. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 1 of *Handbooks in Economics*, Chapter 2, pp. 87–166. Elsevier.

Jäntti, M. and S. P. Jenkins (2015). Income mobility. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of income distribution*, Volume 2 of *Handbooks in Economics*, Chapter 10, pp. 807–935. Amsterdam, Netherlands: North-Holland.

Karush, W. (1939). *Minima of functions of several variables with inequalities as side constraints*. M.Sc. Dissertation, University of Chicago.

Kingma, D. P. and J. Ba (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations*.

Kuhn, H. W. and A. W. Tucker (1951). Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Volume 2, pp. 481–493. University of California Press.

Lange, S., T. Gabel, and M. Riedmiller (2012). Batch reinforcement learning. *Reinforcement Learning 12*, 45–73.

Levine, S., A. Kumar, G. Tucker, and J. Fu (2020). Offline reinforcement learning: Tutorial, review, and perspectives on open problems. https://arxiv.org/abs/2005.01643 (last accessed 01/17/2025).

Li, L., K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar (2018). Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research 18*(185), 1–52.

Lindauer, M. and F. Hutter (2020). Best practices for scientific research on neural architecture search. *Journal of Machine Learning Research 21*(243), 1–18.

Loshchilov, I. and F. Hutter (2019). Decoupled weight decay regularization. In *International Conference on Learning Representations*.

Prince, S. J. D. (2023). *Understanding deep learning*. Cambridge, Massachusetts and London, England: MIT Press.

Roemer, J. E. and A. Trannoy (2015). Equality of opportunity. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 2 of *Handbook of Income Distribution*, Chapter 4, pp. 217–300. Elsevier.

Runje, D. and S. M. Shankaranarayana (2023). Constrained monotonic neural networks. In A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett (Eds.), *Proceedings of the 40th International Conference on Machine Learning*, Volume 202 of *Proceedings of Machine Learning Research*, pp. 29338–29353. PMLR.

Snoek, J., H. Larochelle, and R. P. Adams (2012). Practical Bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems 25*.

Sutton, R. S. and A. Barto (2018). *Reinforcement learning: An introduction* (Second edition ed.). Adaptive computation and machine learning. Cambridge, Massachusetts and London, England: The MIT Press.

A. Numerical details and algorithm

A.1. Bisection search for λ^k and $\tilde{\mathbf{s}}_1^k$

Algorithm 1: Bisection search for λ^k and $\tilde{\mathbf{s}}_1^k$

Inputs:

- Unidimensional cut through social evaluation function: $sef(\mathbf{s}^k) = SEF(\mathbf{s}^k|\mathbf{s}_1^1, \ldots, \mathbf{s}_1^{k-1}, \mathbf{s}_1^{k+1} \ldots, \mathbf{s}_1^K, \mathbf{h}_1^1, \ldots, \mathbf{h}_1^H)$
- Inverse gradient function $f(\nabla_k sef) = \mathbf{s}^k$
- Observed distribution of k^{th} state $\mathbf{s}_1^k$
- Observed gradient $\nabla_k sef = \left[\frac{\partial sef(\mathbf{s}_1^k)}{\partial s_{1,1}^k}, \cdots, \frac{\partial sef(\mathbf{s}_1^k)}{\partial s_{1,N}^k}\right]'$

Initialize variables: $x^{min} \leftarrow \min(\nabla_k sef)$, $x^{max} \leftarrow \max(\nabla_k sef)$;

repeat

  Set $x \leftarrow \frac{x^{min}+x^{max}}{2}$;

  Compute $\hat{\mathbf{s}}^k = f([(x)_{\times N}]')$;

  if $sef(\hat{\mathbf{s}}^k) < sef(\mathbf{s}_1^k)$ **then**

    $x^{max} \leftarrow x$

  else

    $x^{min} \leftarrow x$

until $sef(\hat{\mathbf{s}}^k) \approx sef(\mathbf{s}_1^k)$;

return $\tilde{\mathbf{s}}_1^k = \hat{\mathbf{s}}^k$ and $\lambda^k = x^{-1}$

A.2. Deep Monte Carlo Prediction Algorithm

Algorithm 2: Offline Deep Monte Carlo prediction, estimating $\hat{V}_\pi \approx V_\pi$

Inputs:

- $\mathcal{D} = \{(\mathbf{s_{t,i}}, \mathbf{h_{t,i}}, \mathbf{a_{t,i}}, \mathbf{s_{t+1,i}}, \mathbf{h_{t+1,i}})\}$
- discount factor β
- reward function $u(\cdot)$

Hyperparameters:

- Number of epochs $\#epochs$
- Batch size B
- Number of layers $K+1$ and nodes $\{D_k\}_{k=0}^{K}$
- Type of activation function
- Scale of weight initialization ν
- Optimizer: Learning rate, weight decay, β_1, β_2

Initialize variables: neural network parameters θ_0 of $\hat{V}_\pi(\cdot|\theta_0)$, $k \leftarrow 0$;
Compute returns $U_{t,i} = \left\{\sum_{\tau=t}^{T} \beta^{\tau-1} u(\mathbf{s}_{\tau,i}, \mathbf{h}_{\tau,i}, \mathbf{a}_{\tau,i})\right\}$ for all observations in $\mathcal{D}$;
Transform and normalize input states $\{\mathbf{s_{t,i}}, \mathbf{h_{t,i}}\}$;
for *epoch in* $1, ..., \#epochs$ **do**
  for *batch in* $1, ..., \lfloor \frac{|\mathcal{D}|}{B} \rfloor$ **do**
    $k \leftarrow k + 1$;
    Select random batch $\{(\mathbf{s}, \mathbf{h}, U)\}$ of size B from $\{(\mathbf{s_{t,i}}, \mathbf{h_{t,i}}, U_{t,i})\}$;
    Calculate Loss $L = \frac{1}{B} \sum_{b}^{B} \left(\hat{V}_\pi(\mathbf{s_b}, \mathbf{h_b}|\theta_k) - U_b\right)^2$;
    Optimize model parameters $\theta_{k+1} \leftarrow Optimizer(L, \theta_k)$
  end
end
return $\hat{V}_\pi(\cdot|\theta_k)$;

A.3. Deep Learning Architecture and Hyperparameter Tuning Routine

We train two separate neural networks to estimate the value functions of households with and without young children during the life cycle. For observations in which wives are younger than 39 years, the trajectories are used only to train the relevant network, depending on whether the household has young children. Once wives exceed age 39, the value function is no longer influenced by having young children, so the remaining trajectories can be used to train both networks.

The input to each neural network consists of the assets ($A_{t,i}$), male and female logarithmized wages ($\ln W_{1,t,i}$, $\ln W_{2,t,i}$), and the wife's age t. To reduce the impact of outliers and consequently improve the performance of the neural networks we transform the asset variable by an inverse hyperbolic sine transformation (Burbidge et al. 1988), which – unlike the extended Box-Cox transformation – can cope with zero observations. In addition, we normalize all inputs to the interval $[-1, 1]$. The returns are normalized in the same way. After training, the output of the neural networks is de-normalized to determine the value function at the transformed and normalized grid of Chebyshev knots.

Because the neural networks estimate the value function on a grid of Chebyshev knots that includes areas of the state space not present in the sampled dataset, it is essential to ensure generalizability of the neural network architecture. One approach is to incorporate economic-specific knowledge directly into the pre-activation functions. We enforce monotonicity of the value function by squaring the weights in the pre-activation functions. Hence, the outputs of the $K+1$ layers of the neural network $\mathbf{y} = \mathbf{f}[\mathbf{x}, \phi]$ are

$$
\begin{aligned}
\mathbf{h}_1 &= \mathbf{a}[\beta_0 + (\mathbf{\Omega}_0 \circ \mathbf{\Omega}_0)\mathbf{x}] \\
\mathbf{h}_2 &= \mathbf{a}[\beta_1 + (\mathbf{\Omega}_1 \circ \mathbf{\Omega}_1)\mathbf{h}_1] \\
&\vdots \\
\mathbf{h}_K &= \mathbf{a}[\beta_{K-1} + (\mathbf{\Omega}_{K-1} \circ \mathbf{\Omega}_{K-1})\mathbf{h}_{K-1}] \\
\mathbf{y} &= \mathbf{a}[\beta_K + (\mathbf{\Omega}_K \circ \mathbf{\Omega}_K)\mathbf{h}_K],
\end{aligned}
$$

where $\mathbf{x}$, $\mathbf{y}$, and $\phi = \{\beta_k, \mathbf{\Omega}_k\}_{k=0}^{K}$ are the inputs, the outputs, and the model parameters,

respectively. The bias vectors β_k have dimensions D_{k+1}, which is the number of units in the next layer. All elements of the $(D_{k+1} \times D_k)$ weight matrices $\mathbf{\Omega}_k$ are squared (element-wise), resulting in non-negative weights.[11] Before training we set all biases to zero and initialize the weights by drawing from a truncated normal distribution with mean zero and variance $\sigma^2_{\mathbf{\Omega}_k} = \frac{\nu}{D_k}$, where ν is a tunable scale hyperparameter. The truncated normal distribution is bounded in the interval $(-2\sigma_{\mathbf{\Omega}_k}, 2\sigma_{\mathbf{\Omega}_k})$. The activation function $\mathbf{a}$ is chosen such that it is flexible enough to express all monotonic functional forms. During the hyperparameter search, the bounded exponential linear unit in Runje and Shankaranarayana (2023) performed better than the hyperbolic tangent function – possibly due to the wider non-saturated domain and consequently fewer vanishing gradients, see Figure 4.

A second measure to boost the generalizability of the neural networks is to use the *AdamW* stochastic gradient descent algorithm by Loshchilov and Hutter (2019). In their work, they demonstrate that decoupling weight decay from the gradient-based updates significantly outperforms the original *Adam* optimizer by Kingma and Ba (2015), which relies on L_2 regularization, in terms of both training and test loss on their example datasets.

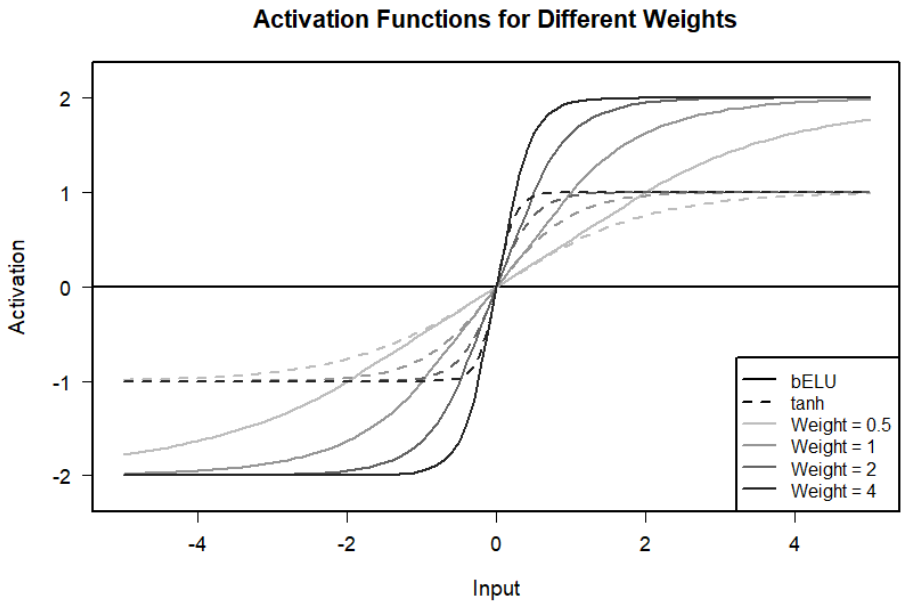
Lastly, the neural networks that result from the hyperparameter search are over-parameterized, which often enhances the interpolation properties, see Prince (2023, pp. 127-132).

To maximize reproducibility, we follow Lindauer and Hutter (2020) and report the details of the hyperparameter search. In a first round of tuning, we explore a wide range of hyperparameter values via the Hyperband algorithm (Li et al. 2018), see Table 7 for the hyperparameter search space in the first iteration of Hyperband. After each of five Hyperband iterations, we narrow the search space by evaluating the best trials. Then, we run 100 trials of the Bayesian Optimization with Gaussian Process algorithm as proposed by Snoek et al. (2012), to arrive at the final hyperparameter configuration reported in the third column of Table 7. The neural network training – i.e., our reduced-form estimates of the value function – is carried out with the deep learning library *keras3* (Kalinowski et al.

[11] In a first step of our hyperparameter search, we also experimented with absolute weights as in Runje and Shankaranarayana (2023). However, squared weights yielded more promising results.

2024). The hyperparameter optimization algorithms are implemented via the *KerasTuner* package (Abdullayev et al. 2023, O'Malley et al. 2019).

Figure 4: Difference between hyperbolic tangent and bounded ELU function



Notes. The figure depicts realizations of the hyperbolic tangent (dashed lines) and bounded ELU function (bELU, solid lines) for various non-negative weights and a bias of zero. The derivative of the activation function with respect to the input is non-zero on a wider domain for the bounded ELU function.

Table 7: Hyperparameters and their configurations

Hyperparameter Name	**Search Space**	**Configuration**
Units	{256, 512, 1024}	512
Layers	$3 \leq$ layers ≤ 7	5
Activation function	{bELU, tanh}	bELU
Learning rate lr	1e-4 $\leq lr \leq$ 1e-2	8e-3
Rate of weight decay per step wd	$0.003 \leq wd \leq 0.005$	0.004
$AdamW$ β_1	$0.85 \leq \beta_1 \leq 0.95$	0.89
$AdamW$ β_2	$0.99 \leq \beta_2 \leq 0.9999$	0.999
ν	{1, 1.5, 2, 2.5, 3, 3.5, 4}	4
Weight operation	$\{(\cdot)^2, \lvert\cdot\rvert\}$	$(\cdot)^2$
Batch size	75	75
Number of epochs	10	10
Validation split	0.2	0.2
Hyperband factor	4	4
Hyperband resources	150 epochs	150 epochs

A.4. Supplementary references

Abdullayev, T. (2023). *kerastuneR: Interface to 'Keras Tuner'*. R package version 0.1.0.6.

Kalinowski, T., J. J. Allaire, and F. Chollet (2024). *keras3: R Interface to 'Keras'*. R package version 1.2.0.

O'Malley, T., E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi, et al. (2019). *KerasTuner*. Python package version 1.4.7.