In Appendix A, we show how to derive a general welfare measure that includes all cohorts that are affected by a tax reform, i.e. current and future generations. Appendix B analyzes the cohort contributions to the aggregate welfare effects of different tax reforms. Appendix C discusses the computational method.

A Generalized welfare calculations

We measure the welfare of future cohort from an ex ante perspective, meaning before any of the labor market shocks have been revealed. Consequently, there will be only one welfare effect per cohort and the relevant state space is

\[ A \times \mathcal{E} \times \mathcal{I} \times \{2, \ldots, J\} \cup \{1, 2, 3, \ldots, \infty\}. \]

The first part of the state space comprises all current generations, i.e. those who have been economically active already in the initial equilibrium and were surprised by a tax reform at some point during their life cycle. The second part contains all future cohorts, i.e. those who enter the economy at some point along the transition path.

A.1 The aggregate welfare function

Now lets look at a partition \( \mathcal{X} \) of this relevant state space. Again \( \mathcal{X} \) is a set of subsets \( X \) of the state space such that \( \bigcup_{X \in \mathcal{X}} X = A \times \mathcal{E} \times \mathcal{I} \times \{2, \ldots, J\} \cup \{1, 2, 3, \ldots, \infty\}. \) For each of the subsets in \( \mathcal{X} \) we define the aggregate welfare function

\[
V(X, T(X), \Psi) = \int_{(a,\eta,i,j) \in X} v_1(a + T(X), \eta, i, j) \phi_1(da \times d\eta \times di \times dj) \\
\quad + \sum_{t \in X} \theta^{t-1} \int_{\mathcal{E} \times \mathcal{I}} v_t(T(X), \eta, i, 1) \phi_t(\{0\} \times d\eta \times di \times \{1\}),
\]

where \( \theta \) is the discount factor for the welfare effects of future generations. Note that this welfare measure is only defined if \( \theta < \frac{1}{1+\rho} \), since otherwise the infinite sum does not converge. For \( \theta = 0 \),
we obviously are in the case of a pure short-run welfare comparison again, since the welfare effects of all future cohorts are given zero weight. Fukushima (2011) shows that in case \( \theta \to \frac{1}{1+n} \), we can interpret our measure as a measure of pure long-run welfare.

We now do the following thought experiment. We want to determine the amount of (hypothetical) transfer we would have to give to group \( X \) in order to make them equally well off along the transition as they were in the initial equilibrium in terms of their aggregate welfare. This means, we want to calculate the transfer \( T(X) \) that solves

\[
V(X, T(X), \Psi_1) = V(X, 0, \Psi_0),
\]

with \( \Psi_0 \) and \( \Psi_1 \) denoting the tax system in the initial equilibrium and along the reform path, respectively.

Using the approximation

\[
V(X, T(X), \Psi) \approx V(X, 0, \Psi) + \frac{T(X)}{\lambda(X)} \left[ \int_{(a,\eta,j,i) \in X} \phi_1(da \times d\eta \times di \times dj) \right. \\
\left. + \sum_{t \in X} \theta^{t-1} \int_{E \times I} \phi_t(\{0\} \times d\eta \times di \times \{1\}) \right]
\]

with

\[
\lambda(X) = \frac{\int_{(a,\eta,j,i) \in X} \phi_1(da \times d\eta \times di \times dj)}{\int_{(a,\eta,j,i) \in X} u_c(c, 1-l) \phi_1(da \times d\eta \times di \times dj)} \\
+ \sum_{t \in X} \theta^{t-1} \int_{E \times I} \phi_t(\{0\} \times d\eta \times di \times \{1\}) \\
+ \sum_{t \in X} \theta^{t-1} \int_{E \times I} u_c(c, 1-l) \phi_t(\{0\} \times d\eta \times di \times \{1\})
\]

we quickly obtain

\[
T(X) = -\lambda(X) \cdot \frac{\int_{(a,\eta,j,i) \in X} \Delta v_t(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj)}{\int_{(a,\eta,j,i) \in X} \phi_1(da \times d\eta \times di \times dj)} \\
\left. + \sum_{t \in X} \theta^{t-1} \int_{E \times I} \Delta v_t(0, \eta, i, 1) \phi_t(\{0\} \times d\eta \times di \times \{1\}) \right]
\]

with \( \Delta v_t(\cdot) = v_t(\cdot) - v_0(\cdot) \).

We can now again aggregate the negative transfers of all groups of the partition. When we use the same discount factor \( \theta \) to discount the transfers of future generations, the aggregate welfare change becomes

\[
\Delta V(X, \Psi) = -\int_{A \times E \times I \times \{2,...,J\}} T(X) \phi_1(da \times d\eta \times di \times dj) \\
\left. - \sum_{t=1}^{\infty} \theta^{t-1} \int_{E \times I} T(X) \phi_t(\{0\} \times d\eta \times di \times \{1\}) \right]
\]

\[
= \int_{A \times E \times I \times \{2,...,J\}} \lambda(X) \cdot \Delta v_t(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj) \\
\left. + \sum_{t=1}^{\infty} \theta^{t-1} \cdot \lambda(X) \cdot \int_{E \times I} \Delta v_t(0, \eta, i, 1) \phi_t(\{0\} \times d\eta \times di \times \{1\}) \right].
\]
A.2 Three examples

We now want to discuss how to partition the state space in order to derive the three welfare functions in the paper.

1. Utilitarian welfare function: For the utilitarian welfare criterion we set
\[ \mathcal{X} = \{ A \times E \times I \times \{2, \ldots, J\} \cup \{1, 2, \ldots, \infty\} \}, \]
i.e. the partition just consists of the whole state space itself. We consequently obtain
\[
\Delta V_{\text{util}}(\Psi) = \lambda_{\text{util}} \cdot \left[ \int_{A \times E \times I \times \{2, \ldots, J\}} \Delta v_1(a, \eta, i, j) \, \phi_1(da \times d\eta \times di \times dj) \\
+ \sum_{l=1}^{\infty} \theta^{l-1} \cdot \int_{E \times I} \Delta v_1(0, \eta, i, 1) \, \phi_l(\{0\} \times d\eta \times di \times \{1\}) \right]
\]
with
\[
\lambda_{\text{util}} = \frac{\int_{A \times E \times I \times \{2, \ldots, J\}} \phi_1(da \times d\eta \times di \times dj)}{\int_{A \times E \times I \times \{2, \ldots, J\}} u_c(c, 1-l) \, \phi_1(da \times d\eta \times di \times dj)}
\]
\[
+ \sum_{l=1}^{\infty} \theta^{l-1} \int_{E \times I} \phi_l(\{0\} \times d\eta \times di \times \{1\}) \cdot \theta^{l-1} \int_{E \times I} u_c(c, 1-l) \, \phi_l(\{0\} \times d\eta \times di \times \{1\}).
\]

2. Cohort based welfare: For the cohort base welfare criterion, we partition the state space cohort by cohort and get
\[
\mathcal{X} = \{ A \times E \times I \times \{j\} \mid j \in \{2, \ldots, J\} \} \cup \{\{t\} \mid t \in \{1, 2, \ldots, \infty\}\}.
\]
The aggregate welfare function therefore becomes
\[
\Delta V^{\text{coh}}(\Psi) = \sum_{j=2}^{J} \lambda^{\text{coh},j} \cdot \int_{A \times E \times I \times \{j\}} \Delta v_1(a, \eta, i, j) \, \phi_1(da \times d\eta \times di)
\]
\[
+ \sum_{l=1}^{\infty} \lambda^{\text{coh},j} \cdot \theta^{l-1} \int_{E \times I} \Delta v_1(0, \eta, i, 1) \, \phi_l(\{0\} \times d\eta \times di \times \{1\})
\]
with
\[
\lambda^{\text{coh},j} = \frac{\int_{A \times E \times I \times \{j\}} \phi_1(da \times d\eta \times di \times dj)}{\int_{A \times E \times I \times \{j\}} u_c(c, 1-l) \, \phi_1(da \times d\eta \times di \times dj)}
\]
and
\[
\lambda^{\text{coh},j} = \frac{\int_{E \times I} \phi_l(\{0\} \times d\eta \times di \times \{1\})}{\int_{E \times I} u_c(c, 1-l) \, \phi_l(\{0\} \times d\eta \times di \times \{1\})}.
\]

3. Individual welfare criterion: Last but not least, the individual welfare partition is
\[
\mathcal{X} = \left\{ (a, \eta, i, j) \mid (a, \eta, i, j) \in A \times E \times I \times \{2, \ldots, J\} \right\} \cup \{\{t\} \mid t \in \{1, 2, \ldots\}\}.
\]
and we get

\[
\Delta V^{\text{ind}}(\Psi) = \int_{A \times E \times I \times \{2, \ldots, J\}} \lambda^{\text{ind,c}}(a, \eta, i, j) \cdot \Delta v_1(a, \eta, i, j) \cdot \phi_1(da \times d\eta \times di) \\
+ \sum_{t=1}^{\infty} \lambda^{\text{ind,f}}(t) \cdot \theta^{t-1} \int_{E \times I} \Delta v_t(0, \eta, i, 1) \cdot \phi_t(\{0\} \times d\eta \times di \times \{1\}).
\]

with

\[
\lambda^{\text{ind,c}}(a, \eta, i, j) = u_c(c, 1 - l)^{-1}.
\]

Note that \(\lambda^{\text{ind,f}}(t) = \lambda^{\text{coh,f}}(t)\), since we always evaluate welfare of future cohorts from an ex ante perspective.
B Welfare weights and cohort welfare contributions

B.1 Short run analysis

In the paper we already showed the welfare weights of different cohorts. We thereby constructed the average welfare weight in the individual measure using the individual weights $\lambda_{\text{ind}}(a, \eta, i, j)$ and the distribution $\phi_1(\cdot)$. Figure 1 shows the average welfare weights by individual wealth holdings for three different generations, i.e.

$$\int_{\mathcal{E} \times \mathcal{I}} \lambda(a, \eta, i, j) \phi_1(\{a\} \times d\eta \times di \times \{j\}).$$

Not surprisingly, the welfare weight increases with individual wealth, since consumption increases with individual assets. Consequently, marginal utility of consumption decreases. The older a cohort,

![Figure 1: Individual welfare weights within generations](image)

...the larger is the difference in weights between the poorer and the richer individuals. The reason for that is that assets become a better proxy for remaining life time consumption when people are older.

Figure 2 shows the cohort contributions to aggregate welfare under the three different welfare measures for the tax reform that maximizes aggregate efficiency, i.e. $\kappa_0 = 0.23$ and $\tau_r = 0.00$. As one can clearly see, the welfare effects of decreasing capital income taxes to zero are almost the inverted effects of the $\tau_r = 0.49$ reform shown in Figure 2 of the paper. Consequently, under the individual welfare measure a large part of the population actually gains from the reform. Only a few young cohorts actually lose.

B.2 Current and future cohorts

Figure 3 reports the cohort welfare effects for all generations for the tax reform that maximizes utilitarian welfare when all generations are taken into account. In contrast to the previous figure, we
standardize the welfare effects by the respective population size, i.e. we report
\[
\frac{\int_{A \times E \times I} \lambda(X) \cdot \Delta \nu_1(a, \eta, i, j) \cdot \phi_1(da \times d\eta \times di \times \{j\})}{C_0 \cdot \int_{A \times E \times I \times \{j\}} \phi_1(da \times d\eta \times di \times \{j\})}
\]
for current and
\[
\frac{\int_{E \times I} \lambda(X) \cdot \Delta \nu_1(0, \eta, i, 1) \cdot \phi_1(\{0\} \times d\eta \times di \times \{1\})}{C_0 \cdot \int_{E \times I} \phi_1(\{0\} \times d\eta \times di \times \{1\})}
\]
for future cohorts. As a consequence, the short-run relative welfare changes of the cohort based and the individual measure are different than in Figure 2 of the paper. The difference with respect to the utilitarian welfare measure is also due to the fact that \(\lambda^\text{util}\) now includes future cohorts. As in Figure 2 of the paper, the short-run welfare effects are again much smaller for the young generations and even more negative for the middle-aged and older cohorts under the cohort based and the individual measure compared to the utilitarian measure. In consequence, both measures report a negative aggregate welfare number for this reform in Table 4 of the paper. Especially for the individual welfare criterion, welfare losses of current generations weigh much stronger than future welfare gains, so that the aggregate welfare effect amounts to a loss in permanent consumption of 2.4%.

Finally, Figure 4 shows the corresponding cohort welfare effects for the tax reform with a zero tax rate on capital income reported in the right column of Table 4 in the paper. As expected, future generations incur a welfare loss from this reform, but current generations gain substantially under the individual welfare measure. These positive effects on current cohorts weigh much stronger than the negative effects in the future, so that aggregate efficiency increases. Under the utilitarian measure, on the other hand, losses are substantial, since it is the cohorts with high marginal utility of consumption that actually suffer from welfare losses.
Figure 3: Welfare effects for each cohort ($\kappa_0 = 0.20, \tau_r = 0.49$)

Figure 4: Welfare effects for each cohort ($\kappa_0 = 0.23, \tau_r = 0.01$)
C Computational appendix

We use two distinct solution algorithms: one to solve the household problem and one to solve for macroeconomic quantities and prices.

C.1 Solving the household problem

We first have to discretize continuous elements of the state space \((a, \eta, i, j)\), respectively the asset dimension. We therefore choose \(\hat{A} = \{\hat{a}^1, \ldots, \hat{a}^{n_A}\}\). We then solve the household problem by backward induction, iterating on the following steps:

1. Compute household decisions at maximum age \(J\) for any \((\hat{a}, \eta, i, j)\). Since households are not allowed to work anymore and they die for sure in the next period, they consume all remaining resources.
2. Find the solution to the household optimization problem for all possible \((\hat{a}, \eta, i, j)\) recursively using a line search method à la Powell, see Press et al. (2001, 406ff.). This algorithm requires a continuous function to optimize. We therefore use an interpolated version of \(v_{t+1}(a, \eta, i, j+1)\). Having computed the data \(v_{t+1}(\hat{a}, \eta, i, j+1)\) at any discrete asset grid point in the last iteration step, we can find a piecewise polynomial function \(s_{p_{t+1,j+1}}\) satisfying the interpolation conditions

\[s_{p_{t+1,j+1}}(\hat{a}^k, \eta, i, j+1) = \int v_{t+1}(\hat{a}^k, \eta', i, j+1)Q(\eta, d\eta')\]  

for all \(k = 1, \ldots, n_A\). We use the multidimensional spline interpolation algorithm described in Habermann and Kindermann (2007).

We choose \(n_A = 25\). We also tried higher values, but the results didn’t change.

C.2 The macroeconomic computational algorithm

We solve for quantities and prices using a Gauss-Seidel procedure in line with Auerbach and Kotlikoff (1987). Starting with a guess for quantities and government policy, we compute prices, optimal household decisions, and value functions. Next we obtain the distribution of households on the state space and new macroeconomic quantities. We then update the initial guesses. These steps are iterated until the initial guesses and the resulting values for quantities, prices and public policy have sufficiently converged.

C.3 Computational efficiency

Our algorithm turns out to be quite efficient. It differs from the one used in Conesa et al. (2009) in that we first apply a more efficient interpolation routine and second allow any choice dimension (consumption, leisure, and assets) to indeed be continuous. Minor differences in simulation results are the consequence. Solving for a long-run equilibrium in the original model of Conesa et al. (2009) takes about 10 to 15 minutes, depending on the calibration.\(^1\) Our simulation approach obtains the same results within 4 to 6 seconds! Computing a complete transition path with 320 transition periods takes about 40 minutes time.

\(^1\) We simulate our models on a regular PC with a Intel® Core™ i7-870 Processor with 2.93 GHz and 8M Cache.
References


