**Online Appendix[[1]](#footnote-1)**

# Required licensing statements

This paper makes use of several licensed datasets that require the following disclaimers.

* The paper employs EU-SILC data (European Commission, Eurostat), longitudinal files (October 2017 distribution). Eurostat has no responsibility for the results and conclusions of this paper.
* The paper employs data from Eurostat, the European Community Household Panel (ECHP) 1994-2001. The responsibility for all conclusions drawn from the data lies entirely with the authors.
* The paper employs the HILDA-CNEF dataset, an equivalized subset of data from the Household, Income and Labour Dynamics in Australia (HILDA) survey provided through the CNEF project. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to FaHCSIA, the Melbourne Institute, or Ohio State University.
* This study has been realized using the data collected by the Swiss Household Panel (SHP), which is based at the Swiss Centre of Expertise in the Social Sciences FORS. The project is financed by the Swiss National Science Foundation.
* The German Socio-Economic Panel Study (SOEP) from the German Institute for Economic Research (DIW), Berlin, were made available through the Cross-National Equivalent File (CNEF), as were the data for Canada (Canadian Survey of Labor and Income Dynamics; SLID).
* U.S. estimates are based on the Current Population Survey (Flood et al. 2015) or the Panel Study of Income Dynamics (PSID, made available via the CNEF).
* The British Household Panel Study (BHPS) was made available through the UK Data Archive. The BHPS is representative only of Great Britain, not the entire United Kingdom.

# Country-year coverage and sources

## Cross-National Equivalence Files (CNEF)

* Great Britain (GBR) 1993-2006 (= British Household Panel Survey, BHPS)
* Germany (DEU) 1985-2017 (= Socio-Economic Panel, SOEP)
* Australia (AUS) 2002-2017 (= Household, Income and Labour Dynamics, HILDA)
* Switzerland (CHE) 2001-2017 (= Swiss Household Panel)
* Canada (CAN) 1994-2010 (= Survey of Labour and Income Dynamics, SLID)
* United States (USA) 1971-1997 (= Panel Study of Income Dynamics, PSID) [select analysis]

## European Community Household Panel (ECHP)

* Austria (AUT) 1995-2000
* Belgium (BEL) 1994-2000
* Denmark (DNK) 1994-2000
* Spain (ESP) 1994-2000
* Finland (FIN) 1996-2000
* France (FRA) 1994-2000
* Greece (GRC) 1994-2000
* Ireland (IRL) 1994-2000
* Italy (ITA) 1994-2000
* Netherlands (NLD) 1994-2000
* Portugal (PRT) 1994-2000

## Survey of Income and Living Conditions (SILC)

* AUT 2004-2014
* BEL 2004-2014
* DNK 2003-2014
* ESP 2004-2014
* FIN 2004-2014
* FRA 2005-2014
* GBR 2007-2015
* GRC 2007-2014
* IRL 2004-2014
* Iceland (ISL) 2004-2014
* ITA 2007-2014
* Luxembourg (LUX) 2004-2014
* NLD 2005-2014
* NOR 2003-2014
* PRT 2007-2014
* Sweden (SWE) 2004-2014

## Current Population Survey (CPS)[[2]](#footnote-2)

* USA 1986-2012 (= Current Population Survey, CPS) [most analyses]

Most analyses pertaining to the US are based on data from the CPS March supplement, an annual survey that asks detailed labor force and income questions. It is a large sample of approximately 70,000 households per year and serves as the source for estimating official poverty rates. The March CPS is not a traditional panel survey, in which a set of respondents are consistently followed over time. Instead, geographic residences are sampled and interviewed on a rotating basis over a period of about a year and a half, regardless of the current occupant. Because the survey is repeated twice in March, however, it is possible to trace a subset of individuals from one year to the next if the individuals are living in the same housing unit in March of both years. A developed literature exists regarding how to match adjacent years of the CPS March Supplement (Feng 2001, 2008; Katz, Tenter, and Sidel 1984; Madrian and Lefgren 2000; Welch 1993). Yet there is no method designed to comprehensively and uniformly produce a complete series of matches over the CPS’s history. Furthermore, as noted by Welch (1993) every study should use its own matching criteria depending on the parameters to be measured. Matching algorithms typically take the form of identifying all matches based on anonymous survey identifiers and then validating or invalidating these “naïve” matches (Madrian and Lefgren, 2000) based on observable characteristics. Because we are interested in measuring income instability, it would be problematic to condition potential matches on characteristics highly associated with income volatility—for example, having persistent labor force status or occupational classification.

Our matching method is motivated by two goals: (1) maximizing potential matches and (2) minimizing any bias introduced by the matching process. The key yardstick for determining whether matches are invalid is the so-called migration flag, which indicates whether someone has moved in the last 12 months—in which case, the individual should not be able to be matched in the CPS. However, while limiting the number of mismatches is important, it is at least as important for our analysis that mismatch rates be consistent, so as not to bias our measurement of trends in income instability. At the same time, the CPS data have undergone many changes over the 1986-2012 period we analyze. In particular, the sample size dramatically increased in the early 2000s to allow researchers to study the expansion of children’s health insurance under the State Children’s Health Insurance Program (SCHIP). For the first 3 years of the SCHIP expansion, household identifiers do not uniquely identify households across surveys. Starting in 2005, an additional household identifier is included in the data, which improves the reliability of matches.

Given the foregoing, we developed a matching method that uses a distance minimization approach to match individuals within households. The household roster in year 1 is joined with the roster in year 2 to produce every possible comparison for individuals from both years. Rather than performing a naïve match based on all available identifiers, in other words, we compare all individuals who share a household ID and find those who are most similar. The distance measure used is a weighted score based on a comparison of state, sex, age, race, the individual’s identifying “line number,” and marital status (in order of highest to lowest weight). Line numbers, which are within-household person identifiers (1-39), and marital status are given the lowest weight, because line numbers are missing in some years and because we do not want to unnecessarily exclude those who experience volatility due to change in family composition. Using this weighted score, we consider two individuals a valid match if both are each other’s closest “neighbor” in the household. We then perform an additional validation procedure using the magnitude of the distance measure to ensure that the match is actually a close one and not just the best possible within the household (for example, a one-person dwelling could match a 30-year old to a 70-year old unless we put some restrictions on the absolute distance allowed between potential matches—in this case, a reasonable age restriction).

The method of matching has several advantages for our purposes. Most important, it produces a large and stable match rate, and a small and stable mismatch rate. Because we do not require that the line numbers match exactly, we are able to reliably match individuals even in years for which line numbers are absent. By considering all possible comparisons, the distance-based approach allows us to find matches without unique person identifiers. Moreover, because we apply this method consistently, we can be certain that any bias created by it is carried through the series and does not affect only years in which line numbers are missing. In addition, the use of distance matching provides a solution to the problem of non-unique household IDs. By comparing everyone across households by household ID, the distance procedure effectively sorts households based on who is most similar. For continuity, we use this method even when the additional household ID becomes available in 2005. Previous research has attempted to match households in the March supplement to the monthly basic file, as the basic file does not include households who are sampled as part of the expansion. We prefer the distance approach because (1) it is unclear that individuals included because of the sample expansion are truly unmatchable across years, and (2) to the extent that we do not trust the household IDs, we do not wish to use them to match to the basic file (Madrian and Lefgren 2000).

# Details on Income Variables[[3]](#footnote-3)

For our analyses, we need to have data on market income (MI) and disposable income (DI). The construction of these variables requires some choices, as detailed next.

## Components of income variables

### Cross-National Equivalence Files (CNEF)

For the data from the CNEF, we employ the constructed variables for household pre-government income [i11101] (which represents the combined income before taxes and government transfers of the head, partner, and other family members) and household post-government income [i11102] (the combined income after taxes and government transfers of the head, partner, and other family members). For Australia, Germany, Switzerland, and the United Kingdom, we have access to the original data sources. For the sake of comparability, however, we rely on the CNEF.

The BHPS sample has been continued as a sub-sample in the Understanding Society/ UK Household Longitudinal Study (UKHLS) panel study. However, changes in the collection of income information lead to a break in the series. We therefore do not use UKHLS data, but rely on SILC data for the years after the last wave of the BHPS.

For analyses that do not require long-term panels, our U.S. data source is the CPS (see above), which allows the construction of one-year panels. However, parts of our analyses use long-term panels, in which case we rely on the PSID, as provided by the CNEF.[[4]](#footnote-4)

### European Community Household Panel (ECHP)

For most countries, income variables in the ECHP are reported net of taxes (but they include taxes in France and Finland). We use variable “HI100 Total net household income” as our measure of DI. It is the sum of the following variables:

* HI111 Wage and salary earnings
* HI112 Self-employment earnings
* HI121 Capital income
* HI122 Property/rental income
* HI123 Private transfers received
* HI131 Unemployment related benefits
* HI132 Old-age/survivors' benefits
* HI133 Family-related allowances
* HI134 Sickness/invalidity benefits
* HI135 Education-related allowances
* HI136 Any other (personal) benefits
* HI137 Social assistance
* HI138 Housing allowance

To derive MI, we take DI, subtract social transfers (variables H131 to HI138), and add taxes (derived from HI020).[[5]](#footnote-5)

### Survey of Income and Living Conditions (SILC)

SILC contains detailed income data. Market income is the sum of cash income from work or property, including employee cash or near-cash income; cash benefits or losses from self-employment”; income from rental of a property or land; and interest, dividends, and profits from capital investments in unincorporated business. To calculate disposable income, we add transfer income. Person-level variables are aggregated to the household level. Transfer income includes unemployment, old-age, survivors’, sickness, and disability benefits; education-related allowances; family or child allowances; periodic payments to people with insufficient resources (referred to as benefits to reduce “social exclusion not elsewhere classified”); housing allowances; regular inter-household cash transfers received; and income received by people younger than 16. We deduct taxes and alimony payments, including regular taxes on wealth, regular inter-household cash transfers paid, and tax on income and social contributions. We do not use the SILC data for Germany, since we take the German data from the CNEF, which provides longer panels. As noted earlier (3.1.1), we use SILC data for the UK/GB only after the end of the original BHPS panel study.

### U.S. Current Population Survey (CPS)

Our measure of market income for the CPS data is based on the CPS’s household gross money income variable. Money income includes earned income (wage and salary income from employment), property and asset income, cash transfer payments (e.g., Social Security, unemployment benefits, and veterans payments), and self-employment income. It also includes lump-sum and one-time payments, such as catch-up payments from Disability Insurance and settlements and distributions from retirement accounts, to the extent that respondents report these as income. Our disposable income variable adds in social insurance and transfer income and subtracts tax liabilities (adding in tax credits, for those whose net tax liability is negative). Taxes are imputed using TAXSIM Version 9 at http://www.nber.org/~taxsim (Bargain et al. 2011; Feenberg and Coutts 1993).

## Cleaning Routines

We work with person-level files (respondents aged 25-60, except for our analysis or income risk over the life cycle) and assign each person his or her household’s (HH) MI and DI income. However, as is common, we adjust HH-level income data by HH size, dividing HH-level incomes by the square root of HH size.

We apply the following standard cleaning routines:

* Drop negative income values.
* Adjust for inflation.
* Bottom-code non-missing income values at each country-year’s weighted p1.
* Top-code non-missing income values at each country-year’s weighted p99.

# Arc percent changes

All results are based on arc-percent changes, not percent changes. Arc-percent changes are calculated as 2\*(Income[t]-Income [t-1])/(Income[t]+Income[t-1]). Unlike percent changes, arc-percent changes are bound by minus and plus 2, and they treat gains and losses symmetrically. For example, a respondent doubling her income from $50 to $100 experiences a 100 percent change (but a 67 arc-percent change). A respondent with a change in income from 100 USD to 50 USD experiences a 50 percent change (but a 67 arc-percent change). The arc-percent approach treats the change of 50 USD in a symmetric fashion.

# Triggers

Based on the small literature on life-events (or “triggers”) that might be associated with income losses (DiPrete and McManus 2000; Gosselin and Zimmerman 2008; Vandecasteele 2010), we focus on a set of triggers that code changes into the following states:

* Unemployment of respondent [pl211X and pl210X]
* Unemployment of HH member other than respondent [pl211X and pl210X]
* Loss of spouse or partner, either through divorce, separation, or death [pb190]
* Sickness of respondent [ph010]
* Sickness of HH member other than respondent [ph010]
* Changes in HH size (increase or decrease) [hx040]

Each of these states is coded as a dichotomous variable, where the value of one indicates that a respondent is in the “bad state” (unemployment, sickness, etc.). With panel data for two consecutive years, four patterns are possible:

* (0 0): no trigger event
* (0 1): transition into bad state = trigger
* (1 0): transition out of bad state
* (1 1): bad state in both time periods

We analyze the association between MI or DI and triggers by regressing the former on the latter, for each trigger separately, and presenting the coefficients as odds ratios (see below).[[6]](#footnote-6) The results in the main text are based on comparisons between (0 1) and (0 0), (1 0), or (1 1). In principle, we could have estimated odds ratios for each of the patterns, but we did not have the space to present these more detailed results. Alternatively, we could have focused on the sample of (0 0) and (0 1) only or restricted the sample in other ways (e.g., by estimating odds ratios for the unemployment trigger on a sample that only includes respondents who were previously full-time employed). This would generally lead to higher odds ratios, as well as higher rates of incidence. However, we not only wanted to keep the sample size as large as possible. We also believe that a natural next step for future research is to estimate models in which several triggers—and their plausible interactions—are included at the same time. Inclusive models of this sort will require the “large sample approach” we have chosen here.

In the main text, we provide a rough estimate of the percentage of market income drops that are associated with a subset of the triggers mentioned above, namely unemployment (self, others in HH), changes in HH size (increase or decrease, loss of spouse/partner, and bad health (self, others in HH)). For each of these triggers, we calculate the percentage of respondents coded as having experienced the trigger, conditional on having experienced a large income drop. Our rough estimate (“40 to 50 percent of all large household income losses occur in conjunction with our measured triggers”) is the sum of these percentages.[[7]](#footnote-7) We stress again that this is only a very rough “back-of-the-envelope” estimate.

One challenge with the analysis of the relationship between events (“triggers”) and income dynamics is that the reference periods for income and events are not always identical. Many variables in the EU-SILC (and most other panel studies) refer to the time of interview (the reference period is “current/at the time of the interview”). Reported income, however, refers to the last calendar year (except in Ireland, where it is the 12 months before the interview; and in the UK, where it is the current calendar year). Fortunately, information on HH size and employment status is available for the income reference period. Subjective health and marital status, however, do not refer to the income reference period. In these cases, we match a respondent’s state with either the previous income reference period, or the current income reference period, depending on whether the interview occurred before or after July. This maximize the overlap between the reference periods in case they are different, although the overlap will still be only partial.

# Robustness

Throughout the paper, we use 25 percent arc income losses as the threshold at which we say a respondent has experienced a “large” income loss. As discussed, there are good reasons for this particular cut-off. Nonetheless, the specific loss threshold chosen does not have much of an effect on our cross-national comparisons, simply because the correlation between indicators based on different cut-off points is very high, as the following table shows.

Table 1: Correlation between of key variables with different cut-off points (10/25/50%)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | 10%+ drop (MI) | 1 |  |  |  |  |  |  |  |
| 2 | 25%+ drop (MI) | 0.94 | 1 |  |  |  |  |  |  |
| 3 | 50%+ drop (MI) | 0.83 | 0.95 | 1 |  |  |  |  |  |
| 4 | 10%+ drop (DI) | 0.90 | 0.85 | 0.76 | 1 |  |  |  |  |
| 5 | 25%+ drop (DI) | 0.85 | 0.86 | 0.82 | 0.95 | 1 |  |  |  |
| 6 | 50%+ drop (DI) | 0.75 | 0.81 | 0.81 | 0.85 | 0.96 | 1 |  |  |
| 7 | Risk reduction (10%) | 0.40 | 0.39 | 0.39 | 0.62 | 0.54 | 0.47 | 1 |  |
| 8 | Risk reduction (25%) | 0.51 | 0.50 | 0.50 | 0.69 | 0.70 | 0.65 | 0.89 | 1 |
| 9 | Risk reduction (50%) | 0.51 | 0.53 | 0.51 | 0.66 | 0.71 | 0.74 | 0.74 | 0.88 |

Note: All correlations are significant at p<=0.01

Not surprisingly, then, the figures we report in the paper (based on 25%+ income losses) generalize to different cut-off points. To illustrate, we reproduce all figures in the paper using a 50%+ cutoff. Though the incidence of losses is obviously lower when a higher threshold is chosen, the relative ranking of countries, over-time trends within nations, and direction and magnitude of the correlation between countries’ loss levels and their social policies are all similar.

# Figures (robustness)

Figure 1: Prevalence of large income losses (50%)



Figure 2: Probability of experiencing at least on 50% arc income drop within time window



Note: Shown are average predicted logistic hazard rates.

**Figure 3: Triggers (50% losses)**



Figure 4: Risk reduction (pre- and post-2008); 50% income losses



Figure 5: Social spending, as well as benefit generosity, and risk reduction (50%)



Note: Generosity indices are from (Scruggs, Jahn, and Kuitto 2014).

Figure 6: Risk and risk reduction over the life-cycle (50% losses)



Note: Shown are average predicted probabilities of experiencing the state indicated on the y-axis, by age. Based on multinomial models regressing the states on quartic age. Y-axes vary.

# References

Bargain, Olivier et al. 2011. “Tax Policy and Income Inequality in the US, 1978-2009: A Decomposition Approach.”

DiPrete, Thomas A., and Patricia A. McManus. 2000. “Family Change, Employment Transitions, and the Welfare State: Household Income Dynamics in the United States and Germany.” *American Sociological Review* 65(3): 343–70.

Feenberg, Daniel, and Elisabeth Coutts. 1993. “An Introduction to the TAXSIM Model.” *Journal of Policy Analysis and management* 12(1): 189–194.

Feng, Shuaizhang. 2001. “The Longitudinal Matching of Current Population Surveys: A Proposed Algorithm.” *Journal of Economic and Social Measurement* 27(1, 2): 71–91.

———. 2008. “Longitudinal Matching of Recent Current Population Surveys: Methods, Non-Matches and Mismatches.” *Journal of Economic and Social Measurement* 33(4): 241–252.

Flood, Sarah, Miriam King, Steven Ruggles, and John Robert Warren. 2015. *Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-Readable Database].* Minneapolis, MN: University of Minnesota [https://cps.ipums.org/cps].

Gosselin, Peter, and Seth Zimmerman. 2008. “Trends in Income Volatility and Risk, 1970-2004.” *Urban Institute Working Paper*.

Hacker, Jacob S. et al. 2011. *The Technical Appendix for the Economic Insecurity and the Great Recession*. http://economicsecurityindex.org/assets/ESI%20Full%20Report%202011.pdf.

Katz, Arnold, K. Tenter, and Philip Sidel. 1984. “Comparison of Alternative Ways of Deriving Panel Data from the Annual Demographic Files of the Current Population Survey.” *Review of Public Data Use* 12(1): 35–44.

Madrian, Brigitte C., and Lars John Lefgren. 2000. “An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents.” *Journal of Economic and Social Measurement* 26(1): 31–62.

Nichols, Austin, and Philipp Rehm. 2014. “Income Risk in 30 Countries.” *Review of Income and Wealth* 60: S98–116.

Scruggs, Lyle, Detlef Jahn, and Kati Kuitto. 2014. *Comparative Welfare Entitlements Dataset 2*. University of Connecticut & University of Greifswald: http://cwed2.org/.

Vandecasteele, Leen. 2010. “Life Course Risks or Cumulative Disadvantage? The Structuring Effect of Social Stratification Determinants and Life Course Events on Poverty Transitions in Europe.” *European Sociological Review* 27(2): 246–263.

Welch, Finis. 1993. “Matching the Current Population Surveys.” *Stata Technical Bulletin* 2(12).

1. The data, replication instructions, and the data’s codebook can be found at <https://doi.org/10.7910/DVN/NZTGVE>. [↑](#footnote-ref-1)
2. This section is heavily based on Hacker et al. (2011). We gratefully acknowledge the work by Austin Nichols and Stuart Craig to assemble the CPS matched files. [↑](#footnote-ref-2)
3. This section partially draws from (Nichols and Rehm 2014). [↑](#footnote-ref-3)
4. In the PSID, gross income includes labor and property income, and excludes employer-provided defined-benefit pension benefits, Social Security and other social insurance payments (unemployment and worker’s compensation), the cash value of means-tested transfers and cash-equivalent in-kind benefits (e.g., food stamps). Gross income does not include tax liabilities. Net income adds in the excluded social insurance and transfer income, and subtracts tax liabilities (adding in tax credits, for those whose net tax liability is negative). Taxes are imputed using TAXSIM Version 9 at <http://www.nber.org/~taxsim> (see also Feenberg and Coutts 1993). Furthermore, we use family income where household income is unavailable. Estimations based on other panel data sources for the United States indicate that the choice of household income or family income measured at the individual level has no substantive impact on estimates. [↑](#footnote-ref-4)
5. Because we rely on a rough indicator of tax rates, the MI data (DI in France and Finland) are likely of lower quality than the DI data in the ECHP. [↑](#footnote-ref-5)
6. We pool across all country-survey-years, which are: AUT (2005-2015), BEL (2005-2015), CHE (2012-2015), DNK (2004-2015), ESP (2005-2015), FIN (2005-2015), FRA (2006-2015), GBR (2006-2015), GRC (2008-2015), IRL (2005-2015), ISL (2005-2015), ITA (2008-2015), LUX (2004-2015), NLD (2006-2015), NOR (2004-2015), PRT (2008-2015), SWE (2005-2015). [↑](#footnote-ref-6)
7. It is beyond this letter to explore what accounts for the remaining income drops. But we speculate that changes in jobs, earnings, and work hours account for some of the income drops. Furthermore, some events/triggers might not persist long enough to be recorded across two waves. Of course, some income drops are due to measurement error. [↑](#footnote-ref-7)