Online Appendix

Why Didn't the College Premium Rise Everywhere? Employment Protection and On-the-Job Investment in Skills

by Matthias Doepke[1] and Ruben Gaetani[2]

October 2023

[1]Northwestern University and LSE, m.doepke@lse.ac.uk.
[2]University of Toronto, ruben.gaetani@utoronto.ca.

A Data Appendix

The main data sources for the empirical analysis and the model calibration are the Panel Study of Income Dynamics for the United States, and the Socio-Economic Panel (SOEP) for Germany. The PSID is conducted by the Survey Research Center (SRC) at the University of Michigan and can be accessed freely via their website.[3] The SOEP is administered by the German Institute for Economic Research, DIW Berlin. More information on the SOEP and how researchers can gain access to it is available on the Institute's website.[4] We employ version 31 of the SOEP data set (2016, doi:10.5684/soep.v31).[5]

Panel Study of Income Dynamics The PSID (Survey Research Center 2016) was conducted on a yearly basis between 1968 and 1996, and every two years from 1997 onwards. The structure of the PSID is a panel in which individuals belonging to a PSID family in 1968 are followed over time as they form a new household or re-join their previous one. We focus on individuals in families belonging to the original SRC sample, which is designed to be representative of the US population.

We focus on male respondents who, at the time of the interview, are identified as their family's head.[6] For consistency with the assumptions on demographics in the model, we restrict attention to individuals aged 25–64 who declare that they only work for someone else. Hence, we exclude self-employed individuals and those who answer that they work for "Both someone else and self."

We define a single education variable for each individual that corresponds to the maximum reported educational attainment, or the maximum number of years of education. Workers are classified as having a high-school degree if they report at least 12 but less than 16 years of education, and as having a college degree if they report at least 16 years of education. We discard individuals with less than 12 years of education. We define real earnings per hour as earnings per hour deflated via the CPI, using 2010 as base year. We discard observations where the resulting real hourly earnings are below 7.50 2010 dollars, or the total number of hours worked is below 500 or above 5,000 in the last year.

[3] https://psidonline.isr.umich.edu/default.aspx

[4] https://www.diw.de/en/soep

[5] To adjust values for inflation, we use the time series of the Consumer Price Index for the United States and for Germany (Organization for Economic Co-operation and Development 2018) retrieved from the Federal Reserve Economic Data website.

[6] The PSID reports that: "Historically, PSID has used the term Head to refer to the husband in a heterosexual married couple and to a single adult of either sex. Starting in 2017, the term "Reference Person" replaced "Head." For more information visit: https://psidonline.isr.umich.edu/Guide/FAQ.aspx.

Potential experience is defined as current age minus (years of education + 6). The employment tenure variable is continuously available starting from 1981. Until 1993, the information is reported as the number of months with the current employer. From 1994 onwards, the information is reported in three separate questions for years, months, and weeks with the current employer, which must be added to obtain a single tenure figure. To maintain consistency with the quantitative model, we then adjust the tenure figure by taking the minimum between the tenure information and $age - 25$. We discard observations for which the resulting months of tenure are more than $12 \times (age - 16)$.

All the statistics in the paper are computed using a system of weights that keeps the age distribution constant in every year to the 1981 distribution.

Socio-Economic Panel The German SOEP (DIW Berlin 2015) has been conducted on a yearly basis since 1984. Similarly to the PSID, the SOEP is a longitudinal study that periodically surveys the same set of families that were interviewed in the original sample. Although there have been expansions in the samples in 1990 (East German sample) and in 1994 (immigrant sample), we focus on the original sample, which only includes families originally from West Germany.

We focus on male respondents aged 25–64 who do not declare themselves to be self-employed. We infer self-employment from the "generated" variable labeled "STIB - Occupational Position."

The SOEP provides separate variables for college attainment and total years of education and training. We define a college degree variable for each individual if college completion is reported at some point in the panel. For those who do not report college completion, we define a unique educational attainment variable for each individual, based on the highest number of years of education recorded in the panel for that individual. We then generate a high-school degree variable if the individual reports at least 10.5 years of education or training. We discard observations that report less than 10.5 years of education or training.

We discard workers who declare less than 20 hours of work per week. We postulate a minimum hourly wage that is equal, in every year, to the 2015 statutory minimum wage (8.50 euros per hour), discounted by the relative price index.

Starting with the system of individual weights provided by the SOEP for the original sample, we construct a system of weights that keeps the age distribution constant over the years to the 1984 distribution, in an analogous manner to the US sample.

B Additional Figures and Tables

	Log of hourly wage (ages 45–54)	
	US, 1981-2013	DE, 1984-2013
	(1)	(2)
Tenure >= 20	0.281	0.122
	(.030)	(.021)
College graduate	0.488	0.475
	(.035)	(.053)
Interaction	-.107	-.166
	(.048)	(.050)
Exper. 3rd degree pol.	Yes	Yes
Year FE	Yes	Yes
# Obs.	7,578	10,137
R^2	0.241	0.268

Notes: Standard errors clustered at individual level in parentheses. For the PSID, the sample consists of male workers aged 45 to 54 who report between 500 and 5,000 hours worked in a year. For the SOEP, the sample consists of male workers aged 45 to 54 who report a minimum of 20 hours worked per week in the last year. Each column consists of a separate regression of log wages on a year fixed effect, a third-degree polynomial of potential experience, and an indicator variable for 20 years of tenure or more on the current job.

Table B.1: Returns to high tenure for college- and less-educated workers in the US (PSID) and Germany (SOEP).

	Dependent variable: Log hourly real wage							
	US (PSID), 1981-2013				Germany (SOEP), 1984-2013			
	High school		College		High school		College	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returns to 10 years of tenure	.129 (.016)	.121 (.012)	.061 (.028)	.002 (.021)	.039 (.021)	.047 (.016)	.052 (.054)	-.053 (.037)
Difference in returns (High school - College)			.068 (.032)	.119 (.024)			-.013 (.058)	.101 (.040)
# Obs.	22,450	21,888	13,422	13,106	28,814	28,416	9,329	9,247
Experience (polynomial)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Standard errors clustered at individual level in parentheses. For the PSID, the sample consists of male workers aged 45 to 54 who report between 500 and 5,000 hours worked in a year. For the SOEP, the sample consists of male workers aged 45 to 54 who report a minimum of 20 hours worked per week in the last year. The dependent variable is hourly log wages. The explanatory variables include years of tenure, years of tenure squared, and an indicator for whether the job started more than one year ago, instrumented using deviations from the average within the employment spell. Regressions include a third-degree polynomial of potential experience.

Table B.2: Returns to tenure using the method in Altonji and Shakotko (1987).

	Model-implied unemployment rate			
	"United States"		"Germany"	
	L-workers	H-workers	L-workers	H-workers
	Panel A: 1980			
All workers	2.52%	2.57%	2.62%	2.67%
Age 25-34	5.93%	5.95%	6.74%	6.77%
Age 35-44	2.04%	2.10%	2.23%	2.27%
Age 45-54	1.13%	1.21%	1.03%	1.11%
Age 55-64	0.97%	1.04%	0.49%	0.54%
	Panel B: 2010			
All workers	2.87%	2.89%	2.99%	3.15%
Age 25-34	5.74%	5.75%	7.43%	7.69%
Age 35-44	2.20%	2.18%	2.42%	2.55%
Age 45-54	1.69%	1.73%	1.41%	1.55%
Age 55-64	1.84%	1.89%	0.71%	0.83%

Table B.3: Model-implied unemployment rate by age and education, computed as the share of workers in each age and education group who are not employed at the end of each model period.

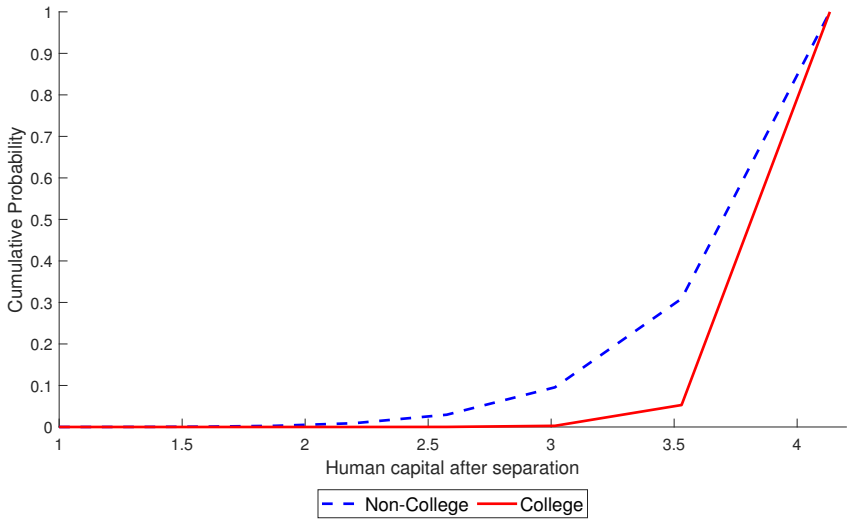

Figure B.1: Cumulative probability of falling to any skill level h' after a separation with skill level $h = h_{10}$ for L (dotted line) and H (solid line) workers.

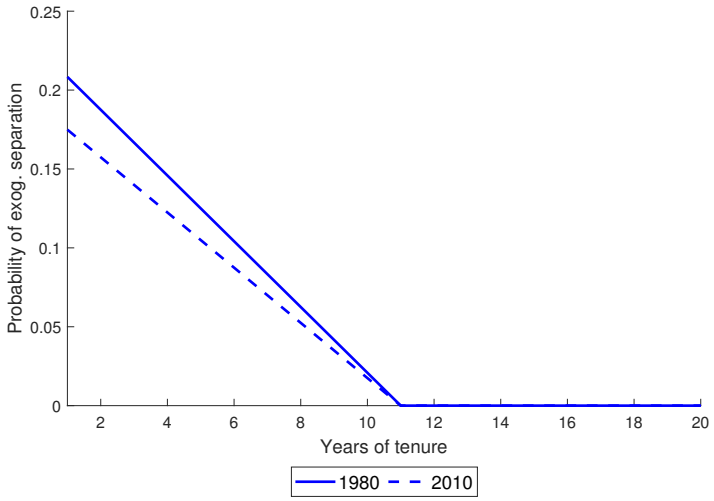

Figure B.2: Probability of exogenous separation as a function of tenure for L and H workers in the 1980 (solid line) and 2010 (dotted line) calibrations.

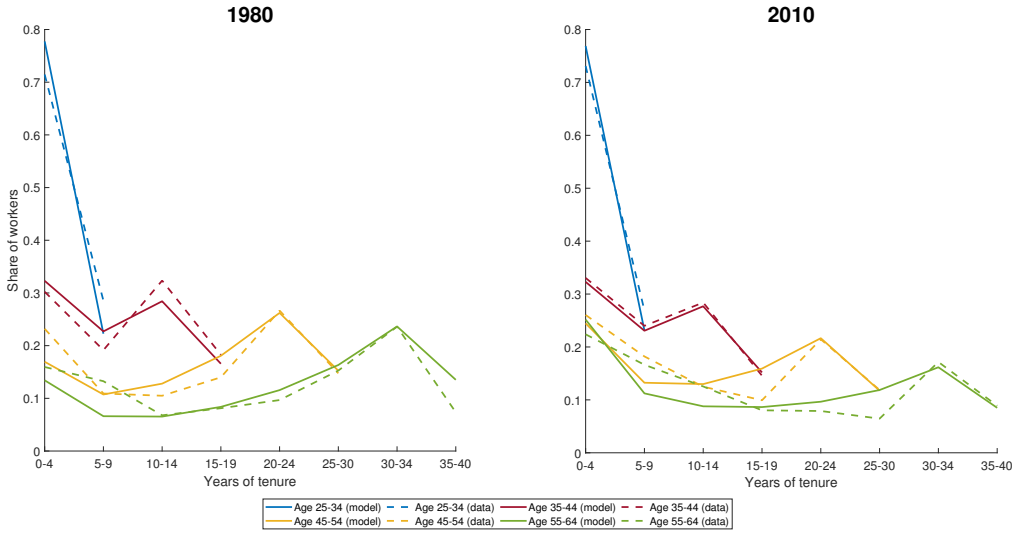

Figure B.3: Distribution of workers by age and years of tenure in the data (dotted lines) and model (solid line), in the 1980 (left panel) and 2010 (right panel) steady states.

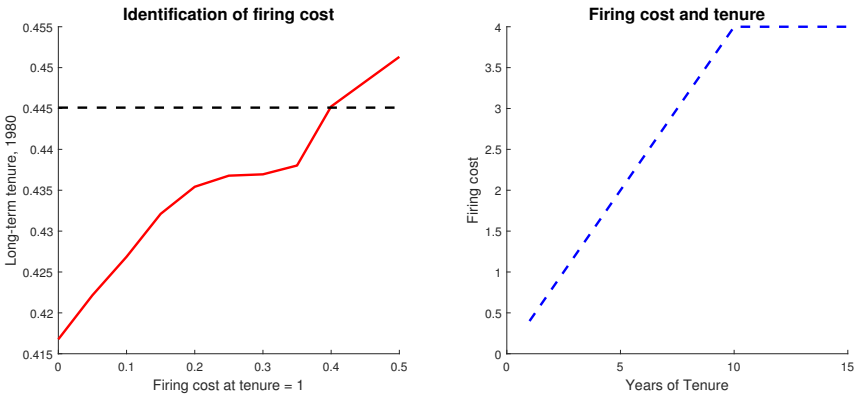

Figure B.4: Left panel: Share of Long-term tenure workers in Germany in 1980 as a function of the firing cost, and share in the data (dotted black line). Right panel: Calibrated firing cost as a function of years of tenure with current employer.

C Model Computation

In computing the model, we impose a vector of 20 possible values for individual human capital, h, log-spaced between $h_1 = 1$ and $h_{20} = 20$. We discretize the vector of turbulence shocks, ϵ, as 10 equally spaced values between 0.01 and $\bar{\epsilon} = 0.6$.

The calibration routine minimizes a loss function defined, for each vector of parameters p, as the sum of the squared differences between the data-generated and the model-generated moment:

$$\min L(p) = \sum_m a_m \left[Data_m - Model_m(p) \right]^2,$$

where the weight a_m is set equal to 10 for the moments that concern the fraction of workers with short- and long-term tenure (which is particularly important to match given our mechanism), and equal to one for the other moments.

D Details on match surplus and wage setting

In this section, we provide detailed expressions for the surplus generated by firm-worker matches, and for the resulting equilibrium wage. To keep notation as simple as possible, we present these expressions for the case of a match not currently experiencing a turbulence shock (the case in which the match is experiencing a turbulence shock is analogous).

The surplus of a match between a firm and a worker of education e, skill level h, and experience x, in a type-p job of tenure t, is equal to:

$$S_p^e(x,h,t) = V_A^e(x,h,t) - E_{h'}\Big[U^e(x,h')\Big] + J_p^e(x,h,t) + \Phi_p^e(x,h,t),$$

where the expectation $E_{h'}$ is taken with respect to the skill loss probabilities $Q^e(i,j)$, as defined in Equations (3) and (4).

If the surplus is positive, the wage is determined by Nash bargaining between the worker and the firm, with the worker retaining a share α of the resulting surplus:

$$V_A^e(x,h,t) - E_{h'}\Big[U^e(x,h')\Big] = \alpha S_p^e(x,h,t).$$

We now have all the ingredients to derive expressions for the equilibrium wage.

In type-A vacancies, the level of effort is not contractible, and is taken as given at the negotiation stage. Using the Bellman equations defined in Section 4.5 and rearranging the expression to isolate the wage, we obtain

$$\begin{aligned} w_A^e(x,h,t) = a^e(x)hz^*(x,h,t)^2 - \beta\Big[&(1-\theta(t))(1-\gamma)E_{h'}\left(V_A^e(x+1,h',t+1)\right) \\ &+ (1-\theta(t))\gamma E_{h',\epsilon}\left(\tilde{V}_A^e(x+1,h',t+1,\epsilon)\right) + \theta(t)E_{h'}\left(U^e(x+1,h')\right)\Big] + \\ &E_{h'}\Big[U^e(x,h')\Big] + \alpha S_A^e(x,h,t). \end{aligned}$$

Analogously, in type-N vacancies the equilibrium wage can be written as:

$$\begin{aligned} w_N^e(x,h,t) = -\beta\Big[&(1-\theta(t))(1-\gamma)V_N^e(x+1,h,t+1) \\ &+ (1-\theta(t))\gamma E_{\epsilon}\Big(\tilde{V}_N^e(x+1,h,t+1,\epsilon)\Big) + \theta(t)E_{h'}\Big(U^e(x+1,h')\Big)\Big] + \\ &E_{h'}\Big[U^e(x,h')\Big] + \alpha S_N^e(x,h,t). \end{aligned}$$

There are three caveats that should be noted. First, the firing cost is only relevant for continuing matches. That is, when the firm and the workers are first matched, the firing cost is not part of the surplus and, if the match is not formed (which never happens in equilibrium) the firm is not subject to the firing cost. Second, skill loss only happens in case of separation after the first period of tenure. That is, if a given match does not result in employment, the worker remains unemployed but is not subject to skill loss. Third, each worker can only be matched with one firm in each period. If the match is not formed, the worker remains unemployed for that period and cannot be matched with another firm.

E Robustness of quantitative results

In this section, we discuss robustness of our main results to alternative choices in our model calibration.

E.1 Magnitude of the turbulence shocks

In the main calibration summarized in Table 4 we set the parameter that controls the minimum productivity loss of regular productivity during turbulent times, $\bar{\epsilon}$, to 0.6. This choice can be interpreted as a normalization, since the model can rationalize the empirical share of workers with long-term tenure with multiple combinations of frequencies of turbulence shocks, γ, and minimum productivity losses, $\bar{\epsilon}$.

To verify that our main results are not driven by this particular choice of $\bar{\epsilon}$, we recalibrate the model by setting $\bar{\epsilon}$ to a lower ($\bar{\epsilon} = 0.4$) and higher ($\bar{\epsilon} = 0.8$) value than our baseline calibration ($\bar{\epsilon} = 0.6$). Columns (1) and (3) of Table E.4 summarize the remaining calibrated parameters in these alternative calibrations. Most parameters are estimated to be very similar regardless of the choice of $\bar{\epsilon}$. Unsurprisingly, the only exceptions are γ_{80} and γ_{10}, with higher values of $\bar{\epsilon}$ implying lower estimated values of the frequency of turbulence shocks.

Table E.5 reports the college wage premium in the 1980 and 2010 steady states and in the counterfactual that isolates the effect of turbulence in the "United States" and "Germany" economies under the baseline calibration (Panel B) and the alternative calibrations using a lower (Panel A) and higher (Panel C) values of $\bar{\epsilon}$. All these versions of the model deliver identical implications of turbulence and labor protection for the college premium.

E.2 Calibrating the model using German SOEP data

The targets of our baseline calibration include, among the other moments, the college premium by age group in the PSID data for the United States. As discussed in Section 6.3, these moments are critical to pin down our mechanism and are key to understand why, as a result of turbulence, the college premium evolves differently in economies with and without employment protection. In particular, a key implication of our mechanism is that the experience-college premium profile should become steeper only in an economy without employment protection. This implication is confirmed by comparing the experience-college premium profile in the PSID and the SOEP data between 1980 and 2010 (Table 9).

Parameter	Interpretation	(1) $\bar{\epsilon} = 0.4$	(2) $\bar{\epsilon} = 0.6$	(3) $\bar{\epsilon} = 0.8$	(4) SOEP
	Panel A: Time-invariant parameters				
σ^L	L skill specificity	0.308	0.309	0.308	(0.309)
σ^H	H skill specificity	0.047	0.053	0.053	(0.053)
ψ	Prob. skill upgrade	0.494	0.498	0.500	0.466
η	Pareto initial skills	4.580	4.574	4.570	4.587
	Panel B: Time-variant parameters (1980)				
g^L_{80}	L productivity growth	0.0005	0.0004	0.0002	-0.0067
A^H_{80}	Skill bias	1.288	1.286	1.282	1.161
γ_{80}	Prob. turbulence shock	0.014	0.020	0.026	(0.020)
$\theta_{80}(1)$	Prob. separation shock	0.209	0.208	0.209	(0.208)
	Panel C: Time-variant parameters (2010)				
Δg_{10}	Change product. growth	0.007	0.007	0.007	0.0001
$\nu^L_{A,10}$	Share of type-A jobs	0.489	0.513	0.538	0.000
$\nu^H_{A,10}$	Share of type-A jobs	0.930	0.950	0.951	0.8541
A^H_{10}	Skill bias	1.495	1.495	1.500	1.235
γ_{10}	Prob. turbulence shock	0.038	0.053	0.068	(0.053)
$\theta_{10}(1)$	Prob. separation shock	0.174	0.175	0.175	(0.175)

Table E.4: Calibrated parameter values for the 1980 and 2010 steady states under different values of $\bar{\epsilon}$ (Columns 1-3) and in the calibration that uses SOEP data (Column 4). Values in brackets are taken from the baseline calibration (Column 2).

We now verify that the same implication holds true when we use the SOEP data to calibrate the model and remove the firing cost to generate the "United States" economy. In performing this alternative calibration, we keep some of the parameters to the baseline calibration and exclude the corresponding targets from the loss function. In particular, we use the magnitude of the turbulence and separation shocks, the transferability parameters, the size of the firing cost, and the parameters controlling the sensitivity of the supply of type-A vacancies obtained in the baseline calibration, since these parameters are only well-identified starting from a setting with no employment protection. We postulate a higher value of workers' productivity during unemployment ($b = 0.4$), to reflect the more generous unemployment insurance prevailing in the German labor market (Ljungqvist

Setting	College Premium	
	No Firing Cost "United States"	Firing Cost "Germany"
Panel A: $\bar{\epsilon} = 0.4$		
1980	0.32	0.32
2010: turbulence and skill-biased tech. change	0.59	0.46
2010: only turbulence	0.43	0.32
Panel B: $\bar{\epsilon} = 0.6$		
1980	0.32	0.32
2010: turbulence and skill-biased tech. change	0.59	0.47
2010: only turbulence	0.43	0.33
Panel C: $\bar{\epsilon} = 0.8$		
1980	0.32	0.31
2010: turbulence and skill-biased tech. change	0.59	0.47
2010: only turbulence	0.43	0.32
Panel D: SOEP Calibration		
1980	0.38	0.37
2010: turbulence and skill-biased tech. change	0.61	0.39
2010: only turbulence	0.53	0.34

Table E.5: The rise in the college wage premium (log difference of average wages) with and without employment protection

and Sargent 2008). We then calibrate the remaining parameters by using SOEP data on the college premium by age group, on the lifetime wage growth of H-workers, and on the dispersion of wages at age 25. Column (4) of Table E.4 shows the resulting calibrated parameters. Starting from this calibration, we then remove the firing cost to generate the "United States" economy in the 1980 and 2010 steady states.

Panel D of Table E.5 shows the college premium in the "United States" and the "Germany" economy in this alternative calibration. The qualitative pattern of the baseline results is preserved. However, as expected, in this case the college premium does not increase significantly in the "Germany" case, consistently with the observed behavior of the college premium in Figure 1. Table E.6 displays the college premium by age group in this alternative calibration. Consistently with our mechanism, we observe a steepening of the

	1980		2010		
	Data	Model	Data	Model	
				Full	Turb.
	Panel A: "United States"				
Age 25-34	0.244	0.174	0.408	0.282	0.174
Age 35-44	0.282	0.314	0.541	0.510	0.421
Age 45-54	0.367	0.440	0.645	0.726	0.655
	Panel B: "Germany"				
Age 25-34	0.157	0.162	0.291	0.219	0.176
Age 35-44	0.308	0.300	0.281	0.341	0.294
Age 45-54	0.413	0.416	0.365	0.423	0.397

Table E.6: College premium by age group in the "United States" (Panel A) and the "Germany" (Panel B) economies in 1980 and 2010 in the calibration using SOEP data.

experience-college premium profile in the "United States" economy (with no employment protection) but not in the "Germany" economy (with employment protection).

References

Altonji, Joseph G., and Robert A. Shakotko. 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies* 54 (3): 437–459.

DIW Berlin. 2015. German Socio-Economic Panel (SOEP), Version 31. https://doi.org/10.5684/soep.v31; Accessed May 2016.

Ljungqvist, Lars, and Thomas J Sargent. 2008. "Two Questions about European Unemployment." *Econometrica* 76 (1): 1–29.

Organization for Economic Co-operation and Development. 2018. "Consumer Price Index: All Items: Total for Germany [DEUCPIALLAINMEI] and Consumer Price Index: All Items: Total for United States [CPALTT01USA661S]." retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org; Accessed February 2018.

Survey Research Center. 2016. Panel Study of Income Dynamics. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. https://psidonline.isr.umich.edu/default.aspx; Accessed April 2016.