Appendix to
The Life-Cycle and the Business-Cycle of Wage Risk - Cross-Country Comparisons

Christian Bayer*
Falko Juessen†
Universität Bonn and IGIER
Technische Universität Dortmund and IZA

This version: January 4, 2011

1 Further details on the estimator

We provide further details on the selection of moments, the weighting matrix employed, and evaluate the properties of the applied estimator in a number of Monte-Carlo experiments. There, we find that the estimator is capable of identifying the cyclicality of wage risk in data sets of a size comparable to the data sets at hand.

1.1 Selection of moment conditions

To apply the estimator, we need to decide on the subset of moments we want to match from the vector of moments $\mathbf{m}$. In their baseline specification, Storesletten et al. (2004) restrict the set of moments to those from a subset of age groups, a theoretical reason for which could be that further moment conditions provide little extra information but only increase estimation uncertainty (i.e. moment conditions being "weak"). However, experimenting with such selected age groups in a set of Monte-Carlo experiments did not show any evidence of a problem of weak moment conditions, see Section 1.3. Correspondingly, increasing the number of moment conditions in general increases the precision of the estimator. Therefore, we use all ages between 30 and 55 and all available years for each sample. We restrict ourselves to this prime-age group in order to eliminate the effect of country-specific differences in household formation (age below 30) and retirement (at ages 55+) that may otherwise influence our results. We assume that households enter the labor market - more specifically, accumulate labor market shocks - at the age of 25 (except for specification V, see Table 1 in the main text).

*IIW, Universität Bonn, Lennéstr. 37, 53113 Bonn, Germany; Tel. +49-228-734073; email: christian.bayer@uni-bonn.de

†Technische Universität Dortmund, Department of Economics, 44221 Dortmund, Germany; phone: +49-231-755-3291; fax: +49-231-755-3069; email: falko.juessen@tu-dortmund.de

1.2 Weighting matrix

The GMM estimator used in this paper employs a weighting matrix to gain efficiency over the simple unweighted minimum distance estimator. In principle, one should weight the moment-distances $(\mu - m)$ with the inverse covariance matrix of each moment condition. In practice, such approach involves estimation of the covariance matrix and can lead to more imprecision in small samples relative to the use of a unitary weighting matrix. We follow Storesletten et al. (2004) and take an intermediate approach, where we estimate the most apparent part of the covariance matrix that results from the definition of our moment conditions themselves. Yet, we ignore the covariance that results from the fact that the sequence of two year panels that we construct from the data partly overlap in the set of households they contain. Instead, we treat the two year panels as if each of them was sampled independently.

This implies that we presume variance ($k = 1$) and autocovariance ($k = 2$) moments, $m_{k,t,h}$ and $m_{k,t+s,h+j}$, are uncorrelated as long as $s \neq 0$ or $j \neq 0$, since then the wage information that they summarize stems from different, independent households. For $s \neq j$, this assumption holds as each household ages by one year each period. For $s = j$, we have to invoke the assumption that the overlap of the two year panels is negligible. As argued, we make this assumption to avoid small sample biases from imprecise covariance estimation.

By contrast, moments $m_{1,t,h}$ and $m_{2,t,h}$ exhibit some clear cut correlation, as they both exploit information of the wages of the same set of individuals at time t (as well as time $t+1$). Moreover, they are also correlated from the fact that the future wage in period $t+1$ of these households that is used to construct $m_{2,t,h}$ is correlated with the wage in period t. The covariance $\boldsymbol{\sigma}_{12}(t,h) = cov(m_{1,t,h}, m_{2,t,h}) \in R^{2\times 2}$ can be easily estimated within sample. The moment conditions refer to the product $\hat{\omega}^2_{i,t,h}$ and $\hat{\omega}_{i,t,h}\hat{\omega}_{i,t+1,h+1}$ at the household level, such that we estimate $\sigma_{12}(t,h)$ as the sample covariance

$$\begin{aligned}\hat{\boldsymbol{\sigma}}_{12}(t,h) = &\frac{1}{N_{th}} \sum_{i=1}^{N_{t,h}} \begin{pmatrix} \hat{\omega}^4_{i,t,h} & \hat{\omega}^3_{i,t,h}\hat{\omega}_{i,t+1,h+1} \\ \hat{\omega}^3_{i,t,h}\hat{\omega}_{i,t+1,h+1} & \hat{\omega}^2_{i,t,h}\hat{\omega}^2_{i,t+1,h+1} \end{pmatrix} \\ &- \frac{1}{N^2_{th}} \sum_{i=1}^{N_{t,h}} \begin{pmatrix} \hat{\omega}^2_{i,t,h} \\ \hat{\omega}_{i,t,h}\hat{\omega}_{i,t+1,h+1} \end{pmatrix} \sum_{i=1}^{N_{t,h}} \begin{pmatrix} \hat{\omega}^2_{i,t,h} & \hat{\omega}_{i,t,h}\hat{\omega}_{i,t+1,h+1} \end{pmatrix},\end{aligned}$$

where N_{th} is the number of observations in each age-year cell. With the correct ordering of moments $\left(\mathbf{m} := \mathbf{vec}\begin{pmatrix}\mathbf{m}_1' \\ \mathbf{m}_2'\end{pmatrix}\right)$ the weighting matrix $\mathbf{W}$ then takes block diagonal form with $N^2_{th}\hat{\boldsymbol{\sigma}}^{-1}_{12}(t,h)$ as diagonal elements. In particular, this means we weight moment conditions more that stem from age-year cells that hold more observations and that are less kurtotic.

1.3 Monte Carlo results for the GMM estimator

To check for potential small sample bias that might be inherent in the GMM method and to understand how the estimator would be affected by reducing the number of moment restrictions by selecting only a subset of age-groups, we run a set of Monte-Carlo experiments. We simulate the data generating process (DGP) for residual wages $\omega_{i,h,t}$ as given in (1) in the main text and use an AR(1) process with autocorrelation ρ_Y to simulate GDP devaitions from trend. We do so for $N = 60$ households for birthcohorts $c = 1, \ldots, 58$ (in line with the birthcohorts in the PSID - 1913 until 1971). We simulate each of these cohorts for $H = 32$ years (again in line with the ages 25-56 we consider). Finally, we restrict the sample to the years $t = c + h = 55, \ldots, 81$. This way we obtain a sample similar to the PSID (though perfectly balanced).

For the DGP we set $\rho = 0.93$, $\sigma_a^2 = 0.04$, $\sigma_\varepsilon^2 = 0.04$, $\left(\bar{\phi} + \underline{\phi}\right)/2 = 0.025$, $\left(\bar{\phi} - \underline{\phi}\right) = -0.045$, and $\rho_Y = 0.6$ roughly in line with the PSID estimates. We assume normally distributed innovations. We compare the estimator that uses all age groups $h = 6, \ldots, 31$ to an estimator that uses only information from $h = 6, 16, 26$ (corresponding to ages 30, 40, 50 in the actual data). Table 1 displays the results of these experiments.

The results show that there is no strong small sample bias on any of the estimated parameters. Yet, average estimates and true parameters differ slightly. The standard errors we obtain and the small sample bias are lower when all moment conditions are imposed, showing that there is no apparent problem from weak moment conditions. The efficiency gain from using *more* moment conditions is most pronounced for the cyclicality parameter $\bar{\phi} - \underline{\phi}$ — which is intuitive as using information on more age groups implies more business cycle informtaion being exploited.

Overall, the standard errors in this experiment are comparable to the ones obtained by running the bootstrap on the actual data as in Table 1 in the main text. Standard errors are slightly smaller in the simulation experiment but in general of the same order of magnitude. The difference is likely due to the fact that the actual data is not balanced in the size of age-year cells as is the experimental data. In particular, we observe that also in the simulated data, the variance of fixed effects cannot be estimated very precisely — i.e. it is only weakly identified.

Table 1: Results from a Monte Carlo experiment for the GMM estimator employed

	ρ	σ_α^2	σ_ε^2	$\left(\bar{\phi}+\underline{\phi}\right)/2$	$\bar{\phi}-\underline{\phi}$
Full Set of Moment Conditions	0.9272 (0.0110)	0.0311 (0.0221)	0.0386 (0.0036)	0.0277 (0.0067)	-0.0452 (0.0068)
Only ages 6,16,26	0.9265 (0.0168)	0.0299 (0.0259)	0.0382 (0.0054)	0.0285 (0.0089)	-0.0450 (0.0123)
True Parameters	0.9300	0.0400	0.0400	0.0250	-0.0450

Results from 5,000 replications of a Monte Carlo experiment of the estimator described in the paper and appendix A using an age-year cell size of $N = 60$. The numbers refer to the average estimate, numbers in brackets give standard deviations.

2 Sample selection

For the PSID, we drop all households that belong to the Survey of Economic Opportunity (SEO) or the Latino sample. For the BHPS, we keep only households living in England.[1] For the GSOEP, we restrict the sample to West German households and drop observations from the migrant and high income sample. Moreover, we remove all households where the household head is a migrant who has immigrated to Germany after the age of ten. We expect the productivity life cycle of migrants to be governed by other institutional factors (e.g. education) than those of non-migrants.[2]

We use the age h of the household head to attach a working and business cycle history to each household in a given year t. We define an age-year (h, t) cell as the sub-sample of households in year t with the same year of birth $c = t - h$, irrespective of the exact date of the survey interview relative to the exact date-of-birth of the interviewee. From the information of the household's labor income y_{it} and hours worked η_{it} we calculate the household's wage rate $w_{it} := y_{it}/\eta_{it}$. We restrict the sample to those households that

[1] The BHPS started with mainly households living in England. In later sample waves, households from Wales, Scotland, and Northern Ireland were added to the BHPS, which implies that these economically diverse parts of the UK do not have a constant sampling weight.

[2] We treat the GSOEP data differently to the other data sources in this respect, because in the GSOEP the year of entry into the country of a household member is easily available. Controlling for the status of a migrant could be important for three reasons: First, it is not clear how comparable years of schooling are for migrants doing their schooling abroad. Second, migrants who enter the country in working age have gone through a different business cycle history. Third, migration often comes in waves that are likely to be related to the business cycle.

Table 2: Number of Observations by Year

Year	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
PSID	1300	1312	1335	1357	1367	1448	1508	1564	1612	1653	1725	1797	1836
BHPS													
GSOEP													

Year	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
PSID	1878	1913	1966	1991	2043	2079	2126	2184	2202	2226	2257	2164	2212
BHPS											1509	1485	1498
GSOEP				1595	1597	1678	1699	1682	1676	1695	1716	1706	1679

Year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
PSID	2616	2692	2588										
BHPS	1487	1538	1570	1735	1621	1702	1685	1492	1485	1486	1432	1415	1389
GSOEP	1670	1662	1641	1608	1877	1809	3428	3191	3085	2963	2814	2663	

The number of observations refers to the number of households in each two-year subpanel with consecutive observations in years $t, t+1$.

Table 3: Sample Selection

	PSID	BHPS	GSOEP
Initial Number of Observations	68219	38733	60488
Eliminated Number of Observations:			
– Due to elimination of outliers, hours, age,			
≥ 2 obs per HH, only HH w/ less than 14 HH members	5269	6734	6995
– Due to non-consecutive observations	7999	7470	8359
Final sample	54951	24529	45134

supply at least 520 hours of market work per year, being equivalent to at least a quarter of full-time employment of one household member (assuming 40 hours per week). In addition, we apply a three standard deviation criterion to the wage rate for each age-year cell of households to identify outliers, which we then remove from the sample. We experimented with alternative outlier selection criteria which in general lead to similar results, see Section 3. Finally, a household enters the sample only if we observe the household at least in two consecutive years since we want to calculate autocovariances. This means we generate a set of two-year overlapping panels from the original data, which we then use for the analysis. Tables 2 and 3 provide information on the number of observations (i.e. households) by year in each two-year panel as well as on the number of observations we loose due to the sample selection.

3 Further robustness checks

Table 4: Additional robustness checks: cyclical indicator and outlier criterion

	GDP-HP(6.25)			Percentile-Outlier-Criterion		
	PSID	BHPS	GSOEP	PSID	BHPS	GSOEP
ρ	.9138	.9275	.8991	.9319	.9352	.9218
	(.0469)	(.0240)	(.0362)	(.0209)	(.0249)	(.0114)
σ_α^2	.0319	.0147	.0441	.0640	.0248	.0253
	(.0484)	(.0242)	(.0178)	(.0389)	(.0269)	(.0079)
σ_ε^2	.0413	.0413	.0319	.0496	.0420	.0316
	(.0085)	(.0041)	(.0037)	(.0059)	(.0044)	(.0018)
$\left(\bar{\phi}+\underline{\phi}\right)/2$	.0306	.0177	.0087	.0213	.0147	.0102
	(.0151)	(.0066)	(.0064)	(.0102)	(.0070)	(.0022)
$\bar{\phi}-\underline{\phi}$	-.0613	-0.0234	.0175	-.0426	.0014	.0204
	(.0218)	(.0091)	(.0049)	(.0202)	(.0088)	(.0044)

See notes to Table 1 in the main text.

The main text provides various robustness checks. Table 4 adds further robustness checks for measuring the cycle and for defining outliers. As one can see, our results are qualitatively and quantitatively robust to the choice of the cyclical indicator (using an HP(λ = 6.25) filter, Ravn and Uhlig (2002)) and to defining outliers based on a percentile criterion instead of standard deviations. For the alternative outlier criterion, we drop only those households that fall in the top-bottom 0.5% percentiles of each age-group, respectively (instead of using a three standard deviation criterion in an age-year cell).

4 Non constant σ_α^2 and σ_ε^2

The specification in the main text restricts σ_α^2 and σ_ε^2 to be constant over time and cohorts. When this restriction does not hold in the data, the estimator cannot match fully the empirical moments $\mathbf{m}(t,h)$, yet it is not straightforward whether this actually generates a bias in the estimation. It will if fluctuations in σ_ε^2 and σ_α^2 are systematically related to the business cycles a cohort experiences throughout their working life.

We can control for time variation in σ_α^2 and σ_ε^2 by assuming that the variance of fixed effects depends on the cohort $c = t - h$ and the variance of transitory shocks is common across cohorts but varies over time. Under these assumptions the formulas for the two moments we generate from the data generalize to

$$\mu_1(t,h) = var\left(\omega_{it}^h | t,h\right) = \sigma_\alpha^2(t-h) + \sigma_\varepsilon^2(t) + \sum_{s=0}^{h-1} \rho^{2s}\phi(Y_{t-s}). \tag{1}$$

$$\mu_2(t,h) = cov\left(\omega_{it}^h, \omega_{it+1}^{h+1} | t,h\right) = \sigma_\alpha^2(t-h) + \rho \sum_{s=0}^{h-1} \rho^{2s}\phi(Y_{t-s}). \tag{2}$$

We can eliminate $\sigma_\alpha^2(t-h)$ and $\sigma_\varepsilon^2(t)$ from the moments that our model tries to match if we subtract both from the empirical and theoretical moments their averages over cohorts for a given point in time and over time periods for a given cohort (akin to a within transformation).

To do so, define

$$\bar{\mu}_1^C(t) = \frac{1}{H}\sum_{h=1}^{H}\mu_1(t,h) = \sigma_\varepsilon^2(t) + \frac{1}{H}\sum_{h=1}^{H}\sigma_\alpha^2(t-h) + \frac{1}{H}\sum_{h=1}^{H}\sum_{s=0}^{h-1}\rho^{2s}\phi(Y_{t-s}) \tag{3}$$

$$\begin{aligned}\bar{\mu}_1^T(c) &= \frac{1}{\tau(c)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\mu_1(t,t-c) \quad (4)\\ &= \sigma_\alpha^2(t-h) + \frac{1}{\tau(c)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sigma_\varepsilon^2(t) + \frac{1}{\tau(c)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sum_{s=0}^{h-1}\rho^{2s}\phi(Y_{t-s})\end{aligned}$$

$$\begin{aligned}\bar{\mu}_2^T(c) &= \frac{1}{\tau(c)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\mu_2(t,t-c) \quad (5)\\ &= \sigma_\alpha^2(t-h) + \frac{1}{\tau(c)}\rho\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sum_{s=0}^{h-1}\rho^{2s}\phi(Y_{t-s})\end{aligned}$$

$$\tau(c) := \min(H+c,T) - \max(1+c,1) + 1 \tag{6}$$

Now we define moments to be matched as

$$\begin{aligned}\mu_1(t,h) - \bar{\mu}_1^C(t) - \bar{\mu}_1^T(t-h) = &\sum_{s=0}^{h-1}\rho^{2s}\phi(Y_{t-s})\\ &- \left[\frac{1}{H}\sum_{h=1}^{H}\sigma_\alpha^2(t-h) + \frac{1}{H}\sum_{h=1}^{H}\sum_{s=0}^{h-1}\rho^{2s}\phi(Y_{t-s})\right]\\ &- \left[\frac{1}{\tau(t-h)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sigma_\varepsilon^2(t) + \frac{1}{\tau(t-h)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sum_{s=0}^{h-1}\rho^{2s}\phi(Y_{t-s})\right]\end{aligned}$$

and analogously for $\mu_2(t,h) - \bar{\mu}_1^T(t-h)$. Let $\bar{\sigma}_\alpha^2$ be the expected value of $\sigma_\alpha^2(c)$ and $\bar{\sigma}_\varepsilon^2$ be the expected value of $\sigma_\varepsilon^2(t)$ then approximately $\frac{1}{H}\sum_{h=1}^{H}\sigma_\alpha^2(t-h) = \bar{\sigma}_\alpha^2$ for all t and

Table 5: Non-constant variances of fixed effects and transitory shocks

	Non-constant σ^2_α and σ^2_ε		
	PSID	BHPS	GSOEP
ρ	0.9688	0.9580	0.9071
	(0.0395)	(0.0378)	(0.0150)
σ^2_α	0.0108	0.0833	0.0306
	(0.0561)	(0.0308)	(0.0128)
σ^2_ε	0.1472	0.0075	0.0002
	(0.0476)	(0.0292)	(0.0120)
$\left(\bar{\phi}+\underline{\phi}\right)/2$	0.0099	0.0088	0.0163
	(0.0103)	(0.0102)	(0.0032)
$\bar{\phi}-\underline{\phi}$	-0.0197	0.0048	0.0327
	(0.0207)	(0.0096)	(0.0064)

See notes to Table 1 in the main text.

$\frac{1}{\tau(t-h)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sigma^2_\varepsilon(t)=\bar{\sigma}^2_\varepsilon$ for all $c=t-h$. Under these approximations

$$\begin{aligned}\mu_1(t,h)-\bar{\mu}_1^C(t)-\bar{\mu}_1^T(t-h)\approx&\sum_{s=0}^{h-1}\rho^{2s}\phi\left(Y_{t-s}\right)\\&-\left[\bar{\sigma}^2_\alpha+\frac{1}{H}\sum_{h=1}^{H}\sum_{s=0}^{h-1}\rho^{2s}\phi\left(Y_{t-s}\right)\right]\\&-\left[\bar{\sigma}^2_\varepsilon+\frac{1}{\tau(t-h)}\sum_{t=\max(1+c,1)}^{\min(H+c,T)}\sum_{s=0}^{h-1}\rho^{2s}\phi\left(Y_{t-s}\right)\right],\end{aligned}$$

from which we can estimate $\left(\rho,\bar{\phi},\underline{\phi},\bar{\sigma}^2_\varepsilon,\bar{\sigma}^2_\alpha\right)$ using a similar GMM estimator as before. Table 5 displays the estimation results for this specification. The results are qualitatively similar to our baseline specification. In particular the results regarding cyclicality do not change substantially. Yet, the estimator looses efficiency and hence some estimates are no longer significant.