A Bayesian approach to labour market modelling in dynamic microsimulation

3. december 2013

Hans Bækgaard
hab@dreammodel.dk
+45 33 30 79 60

Abstract

The paper presents a novel approach to modelling labour market processes in dynamic microsimulation. The method combines and integrates Bayesian simulation based estimation and simulation of the dependent variables. The approach is applied to a dynamic panel model for hourly wage rates for Danish employees using a large panel data set with 17 years of data for 1995 to 2011. The wage rate model and a parallel model for annual work hours are currently being implemented in SMILE (Simulation Model for Individual Lifecycle Evaluation), a new dynamic microsimulation model for the Danish household sector.

The application benefits from the richness of Danish administrative panel data. Nevertheless, the results and the approach have several features that should be of interest to micro-simulators and others. Indeed, the model features both an extraordinarily comprehensive list of dependencies and a rich dynamic structure. Together, these features contribute to ensure that simulations produce realistic cross-sectional distributions and interactions as well as inter-temporal mobility – the key determinants of the quality of a dynamic microsimulation model. In addition to the ‘usual’ socio-demographic variables (gender, age, ethnicity, experience, education etc.), the dependencies include a more novel set of variables that represent a person’s labour market history, secondary school grade and social heritage (represented by the parent’s education level). The dynamic model structure includes a lagged dependent variable, an auto-correlated error term with a mixed Gaussian distribution for the white noise component, an individual random effect with a mixed Gaussian distribution and permanent effect of a person’s first wage after leaving the education system. The estimation sample is identical to the simulation sample, which allows us to use the same historical detail as well as estimated individual effects – i.e. random effect components – for the simulation of future wage rates.

The Bayesian estimation method handles missing observations for the dependent variable – due to either non-employment or temporary non-participation – by treating missing observations as latent variables that are simulated alongside the Bayesian iterations. As a byproduct, the estimations produce model consistent latent wage rates for the unemployed that are useful for labour supply analysis.

1 The work presented in this paper builds upon previous work documented in Bækgaard (2010) that has been extending and elaborated for the purpose of adaptation to a dynamic microsimulation.
1 Introduction

The paper presents the results from a dynamic panel model for hourly wage rates for a panel of Danish employees from 1995 to 2010. A Bayesian technique is used for estimating the dynamic panel model that includes a novel approach for handling missing observations for the dependent variable. This is achieved by treating the missing observations as latent variables that are simulated alongside the model's parameters during the Bayesian iterations. As a bi-product, model consistent values for the missing dependent variables are simulated.

The method is applied to hourly wage rates in the Danish Salary Register (Lønregistret) where wages rates are absent in years when a person is either out of employment or the employer is not participating because the workplace is not in scope due to small size. The analysis provides new interesting insights into the dynamics of the wage process as well as an understanding of the factors that drive hourly wage rates. The estimated equations are used for forecasting the distribution of wage rates in the new Danish dynamic microsimulation model SMILE.

The analysis takes advantage of the Danish administrative registers with comprehensive demographic and labour market related information, including detailed histories of social transfer receipts, secondary school grade and social heritage (represented by the parent's education level). The social security history provides important information about the duration of periods out of employment and how they impact on subsequent income earning potential. The effect of unemployment spells on future and hence potential wage rates is estimated directly through the inclusion of variables for the duration of periods out of work. Importantly, this represents an alternative and effective way to account for the well-known Heckman selection bias (Heckman, 1979). The results are interesting in their own right and demonstrate that unemployment and other periods of non-employment have a significant negative impact on expected subsequent wages and the negative effect increases with the duration of periods out of work.

The dynamic model structure includes a lagged dependent variable, an auto-correlated error term with a mixed Gaussian distribution for the white noise component, an individual random effect with a mixed Gaussian distribution and permanent effect of a person's first wage after leaving the education system. The estimation sample is identical to the simulation sample, which allows us to use the same historical detail as well as estimated individual effects – i.e. random effect components – for the simulation of future wage rates.

The combination of a rich background information and extensive use of individual historical wage formation processes is an unusual but effective way of ensuring realistic simulation of cross-sectional and inter-temporal distributions. The mixed Gaussian distributions for the residuals and the random effect illustrate the importance of taking non-normality into account.

---

2 The wage rate model’s specification has much in common with the model in Geweke and Keane (2000).
2 The SMILE labour market briefly

The wage rate model is an integrated component of SMILE’s labour market module that simulates transitions between labour market statuses and earned income for employed people. The labour market module has the following main components:

*Labour market transitions*

The transitions between labour market statuses are governed by a combination of survival models, annual transitions and rule-based allocations. The model is hierarchical and the transitions are executed in two steps. The upper level determines the main labour market status defined by the four distinct categories: employed, unemployed, temporary out of the labour force, and permanent out of the labour force (Bækgaard, 2014a, forthcoming). The transitions between employment, unemployment and temporary out of the labour force are based on hazard functions while permanent exits from the labour force are determined by annual retirement events.³

The lower level labour market categories provide more detail such as distinctions between wage earners and self-employed and type of benefit entitlements and participation in Government labour market programmes for the non-employed.

*Wage rates and potential wage rates*

This is the topic of the present paper.

*Annual (full-year) paid work hours for wage earners*

Annual worked hours are simulated by a dynamic panel model, which has much in common with the wage rate model (Bækgaard, 2014b, forthcoming). It is an annual model based on annualized paid work hours, which represent the full-year number of hours a person is paid for if the person is employed all months of the year. Persons who are employed part of the year thus have their worked hours scaled down from full-year to the number of months they are actually employed.

Persons who have gaps in the estimation data either because they have not worked at all during the year or did not have employment covered by the Danish Salary Register (Lønregisteret) will be treated as missing observations and treated in the same way as missing observations in the wage rate model.⁴

The dependent variable is annual paid hours as opposed to annual actual hours worked, which is the concept adopted by national accounts in most countries as well as by the OECD. This is important, as it allows us to derive annual earned income as a product of number of months employed, annualized paid hours and hourly wage rates.

*Annual employment income for the self-employed*

Income from self-employment is simulated by a dynamic panel model for annual taxable income and based upon tax return data. The model structure and general approach is otherwise similar to the wage rate model for employees.

---

³ See for example Arnberg and Stephensen (2013).
⁴ The worked hours and wage rate models are both based on data from the Danish Salary Register.
3 The wage rate model

The model’s empirical specification is a general dynamic panel model with multiple sources of heterogeneity and inter-temporal dependencies. The initial period is represented by a separate non-dynamic wage rate equation (1) while equation for the remaining periods (2) include dynamics through a lagged dependent, a first-order autoregressive error term (3) and (4), a random effect (5) and first-period error term contingency (6):

(1) \[ y_{i1} = x_{i1}\beta_1 + \varepsilon_{i1} \]

(2) \[ y_{it} = \gamma y_{i,t-1} + x_{it}\beta_2 + \tau_i + \phi \varepsilon_{i1} + \varepsilon_{it} \quad (t>1) \]

(3) \[ \varepsilon_{i2} = \rho_1 \varepsilon_{i1} + e_{i2} \]

(4) \[ \varepsilon_{it} = \rho_2 \varepsilon_{i,t-1} + e_{it} \quad (t>1) \]

(5) \[ \phi = \sum_{g=1}^{2} e_{n}^{g} \phi_{g} \]

(6) \[ \tau_i = \sum_{j=1}^{m} e_{ij}^{g} (\alpha_{j} + \sigma_{j} \xi_{j}) \quad \xi_{j} \sim n(0,1) \text{ and } P(e_{ij}^{g} = 1) = p_{j}^{g} \]

(7) \[ \varepsilon_{i1} = \sum_{g=1}^{2} \sum_{j=1}^{m} e_{ij}^{g} (\alpha_{j1}^{1} + \sigma_{j}^{1} \xi_{j}) \quad \xi_{j} \sim n(0,1) \text{ and } P(e_{ij}^{g} = 1) = p_{j}^{g} \]

(8) \[ e_{ij} = \sum_{j=1}^{m} e_{ij}^{g} (\alpha_{j2}^{2} + \sigma_{j}^{2} \xi_{j}) \quad \xi_{j} \sim n(0,1) \text{ and } P(e_{ij}^{g} = 1) = p_{j}^{g} \quad (t>1) \]

\( y_{it} \) is the logarithm of the hourly wage rate for person \( i \) in year \( t \), \((x_{i1}, x_{it})\) are the background variables for the first year \((t=1)\) and the subsequent years \((t>1)\) with parameters \( \beta_1 \) and \( \beta_2 \) respectively. \( \tau_i \) is an individual random effect. \( \gamma \) is the parameter for the lagged dependent. \( \rho_1 \) and \( \rho_2 \) are the AR parameters for period 1 to 2 and for subsequent periods respectively. The parameter \( \phi = (\phi_1, \phi_2) \) is a permanent effect of the first period’s error. The effect differs for persons who are entering the estimation panel because they have obtained employment for the first time after leaving the education system \((\phi_1)\) and persons who are entering the estimation panel at a later stage in their labour market career \((\phi_2)\). Hence \( \phi_1 \) represents the persistent effect of the wage rate in a person’s first job on the wage rate in subsequent years. The distinction between education leavers and others in the first year has the added implication that the variance of first-period error term is group specific.

The somewhat tedious notation in (6), (7) and (8) represents a mixed Gaussian distribution for the random effect and the error terms. A mixed Gaussian (or normal) distribution is a combination of two or more – in our case a maximum of four – normal distributions that, when combined, emulate the empirical error term distributions and, as we shall see, it often does so far better than a single normal distribution.
Using the random effect to illustrate, assume that \( \{ \tau_i \}_{i=1,\ldots,n} \) is an observed sample from a mixed normal distribution of \( m \) normal distributions \( n(\mu_j, \sigma_j^2) \), that is

\[
\tau_i = \sum_{j=1}^{m} e_{ij} (\alpha_j + \sigma_j \xi_j) \quad \xi_j \sim n(0,1)
\]

\( e = \{ e_{ij} \}_{i=1,\ldots,n; j=1,\ldots,m} \) is an indicator matrix for which of the \( m \) normal distributions and individual belongs to and \( P(e_{ij} = 1) = p_j \) its associated probability.

The model was estimated by a Bayesian estimation method where missing observations for the dependent variable are treated as latent variables and simulated iteratively alongside the model’s parameters. Appendix 2 provides a detailed description of the estimation procedure, including the specification of the prior and posterior distributions as well as the method for simulating the missing observations.

The parameters of the mixed Gaussian error term and random effect distributions are estimated with a Bayesian method, which is integrated with and applied iteratively with the estimation of the model’s other parameters (see Appendix 2). However, the final parameter values for the mixed Gaussian distributions reported here were estimated ex post using residuals calculated from the final parameter values in (1) to (5). These final parameter values were estimated with a different method, namely the cluster technique in the R-application MCLUST (see Fraley et al, 2012). It is a general method of fitting a mixed normal of several distributions to a sample of values – in our case, a sample of residuals from a panel model.

The model has been estimated separately for the 12 groups defined by gender and the six main educational groups. The background variables encompass a host of personal and labour market relevant characteristics such as age, marital status, number and age of children, experience, secondary school results, type of education and parents’ education. A person’s labour market history is represented by the number of weeks a person has received income-replacing transfers of different types in the current year and during the previous three years.
4 Data

The estimation data are a panel data set that, in principle, covers the full Danish population of 18 to 67-year old persons from 1995 to 2010, the years for which useful wage rate information is available from the Danish Salary Register (Lønregisteret). However, only a subset of the population could be used for estimation purposes. This is in part because compulsory reporting to the Salary Register does not apply to private employers with less than 10 employees, but also because the full data set is too large to be dealt with computationally given the model’s size and the resulting demands on computer memory capacity.

The dependent variable is the wage rate calculated as the average wage (including pension contribution) per paid work hour for all jobs held during the year that are included in the Salary Register. The analysis is restricted to 18 to 67 year olds. Graduation restarts the clock in the sense that a person starts a new period in the year of completing an education.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of persons and observations</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Males</strong></td>
</tr>
<tr>
<td>Primary school</td>
</tr>
<tr>
<td>Secondary school</td>
</tr>
<tr>
<td>Vocational</td>
</tr>
<tr>
<td>Short tertiary</td>
</tr>
<tr>
<td>Medium tertiary</td>
</tr>
<tr>
<td>Long tertiary</td>
</tr>
<tr>
<td>All males</td>
</tr>
</tbody>
</table>

| **Females** | | | | | | | |
| Primary school | 401,880 | 8.3 | 33,490 | 1,306 | 32,184 | 224,453 | 257,943 |
| Secondary school | 129,591 | 33.3 | 43,197 | 11,508 | 31,689 | 186,054 | 229,251 |
| Vocational | 494,424 | 5.6 | 27,468 | 4,572 | 22,896 | 221,148 | 248,616 |
| Short tertiary | 68,331 | 33.3 | 22,777 | 4,572 | 18,205 | 167,109 | 189,886 |
| Medium tertiary | 330,510 | 6.7 | 22,034 | 6,777 | 15,257 | 174,058 | 196,092 |
| Long tertiary | 107,883 | 30.0 | 32,365 | 11,266 | 21,099 | 208,005 | 240,370 |
| All females | 1,532,619 | 11.8 | 181,331 | 40,001 | 141,330 | 1,180,827 | 1,362,158 |

Source: The Danish Salary Register, Denmark’s Statistics (own calculations).

The Salary Register is based on compulsory reporting by employers though not all employers are covered. Apart from non-compulsory participation for private sector workplaces with less than 10 employees, the sample is affected by improved attrition rates over the initial years of the sample period. The compound result is a sample of 1,388,625 males and 1,362,158 females (cf. table 1). On average, 12.8 per cent of the males and 11.8 per cent of the females were used. As a result, the estimation panel represents 169,850 males (1,388,625 annual observations) and 181,331 females (1,362,158 annual
observations) with each person being in the sample on average 8.2 years for males and 7.5 years for females. Missing wage rate observations due to either non-employment or non-coverage is a substantial issue. A higher representation for females is mainly a result of more females working in the public sector and more males than females working at small out-of-scope work places.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Average hourly wage rates for males and females by education in 1995 and 2010 (2010 DKR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>DKR</td>
<td>DKR</td>
</tr>
<tr>
<td>Primary school</td>
<td>159</td>
</tr>
<tr>
<td>Secondary school</td>
<td>200</td>
</tr>
<tr>
<td>Vocational</td>
<td>179</td>
</tr>
<tr>
<td>Short tertiary</td>
<td>195</td>
</tr>
<tr>
<td>Medium tertiary</td>
<td>232</td>
</tr>
<tr>
<td>Long tertiary</td>
<td>280</td>
</tr>
<tr>
<td>All persons</td>
<td>192</td>
</tr>
</tbody>
</table>

Source: The Danish Salary Register, Denmark’s Statistics (own calculations).

Average wage rates differ across the 12 estimation groups defined by gender and educational attainment (table 2 and figures 1a and 1b). As expected, there is generally a positive relationship between the average wage rate and education length. Males have higher average wages than females for all education groups. In general, there is a clear positive relationship between the average wage rate and education length and the gain from taking an education is larger for males than for females. This is particularly the case for medium length tertiary educations.

Source: The Danish Salary Register, Denmark’s Statistics (own calculations).
The background variables such as age, gender and education originate from various administrative registers. The benefit histories are from the Ministry of Employment’s DREAM database with detailed week by week recipient information. All background variables have full coverage – including the observation (persons and years) where no wage information is available.
5 Simulating wage rates

The manner in which the wage rate model is incorporated into the dynamic microsimulation SMILE relies heavily on the fact that the estimation sample is identical to SMILE’s base population. Indeed, this allows us to correctly evaluate the individual specific factors in equations (1) to (8) prior to simulation start.

The simulation of future wage rates is performed as an extrapolation of the historical information obtained by the Bayesian estimation where the estimated parameters from (1) to (8) are used to evaluate the residuals and the random effect parameters for the full working aged population in SMILE’s base data. The wage rate model provides potential wage rates for everyone regardless of whether a person is actually working or not.

More specifically, the following steps are followed as part of creating the necessary information for SMILE’s starting population:

1. Evaluate initial period residuals \( \hat{e}_{it} = y_{it} - x_{it}\beta_i \) for the first year a person appears in the panel with a wage rate.
2. In the initial year of the data set, identify if a person is an education leaver \( (\phi^1) \) or not \( (\phi^2) \)
3. Evaluate residuals for subsequent years from (2), (3) and (4)
4. Evaluate the distribution of \( \tau_i \) using its conditional posterior distribution (see Appendix 2)
   \[
   \tau_i \sim N((X^5'X^5)^{-1}X^5'Y^5, (X^5'X^5)^{-1})
   \]
   The \( \tau_i \)-values are drawn from this distribution.
5. For persons who are present in the estimation panel, but not in the final year (the initial year of the simulation) either due to non-employment or non-participation, the wage rate is updated to that year by applying the wage rate model in a forecasting mode.
6. For persons who are not in the estimation panel at all and persons who reach working age during simulations, the wage rates are simulated with (1) to (8) by drawing residuals from their estimated distributions.

Steps 3. and 4. are complicated by the fact that individuals may have intermediate years with missing information that is, they have gaps in the estimation panel with no wage rate information. For this reason, these steps are performed by applying the estimation technique described in Appendix 2 and 3, but without the random draws from the posterior distributions for the model’s non-individual specific parameters \( (\rho_1, \rho_2, \gamma, \beta_1, \beta_2, \phi, \sigma^2, \sigma^2_i, \sigma^2_2, \sigma^2_\gamma) \) replaced by the estimated parameters (see Appendix 1).
6 Estimation results

This section provides an account of the results of the estimation of the model described in Section 3. The detailed parameter results are shown in Appendix 1. The objective of the following outline is to provide an interpretation of key messages from the parameter results.

The model has been estimated separately for the twelve groups defined by gender and the six main groups of highest education namely primary school, secondary school, vocational, short tertiary, medium tertiary and long tertiary. This approach holds both practical and analytical advantages compared with a joined estimation. Firstly, the inter-temporal parameters \( (\gamma, \rho_1, \rho_2, \phi, \sigma^2) \) and the error term variation \( (\sigma^2_{11}, \sigma^2_{12}, \sigma^2_{22}) \) differ substantially across the groups. Accounting for these differences is more manageable with separate estimations. This is important for consistency of estimation as well as for accuracy of the simulated missing wages. Secondly, separate estimations provide group specific estimates for all the model parameters and these differences provide new and interesting insight into the wage process.

The model is estimated with wage rates that have been indexed to 2010-level for each group so as to avoid heteroscedasticity caused by wage inflation.

The impact of non-employment on wages

The impact of unemployment and other types of non-employment on wage rates is represented by separate variables for the number of weeks on unemployment benefits, social security (unemployed and not unemployed), sickness benefits, salary subsidised employment, maternity leave benefits, labour market leave benefits and other benefits for persons that are not in the labour force. We focus on the duration of benefits during the three years prior to the current calendar year measured as the share if the total number of weeks with benefits.

The parameter estimates represent a (generally) negative wage rate effect of periods of non-employment. A parameter estimate of -0.15 (figure 2a) indicates that one year of unemployment

---

5 The \( \beta_1 \) - and the \( \beta_2 \) -parameters in (1) and (2) are shown in table A1 and A4 respectively. To facilitate comparison across groups and with the \( \beta_1 \) -parameter estimates, the \( \beta_2 \) -parameters have been scaled up by group specific estimates for \( 1/(1-\gamma) \), cf. table A4. Table A2 shows estimates for the dynamic parameters and Table A3 the estimated parameters for the mixed Gaussian residuals and random effects.

6 The benefit history variables are potentially endogenous and, as a result, we cannot say with certainty, that the negative parameters represent a causal effect. Indeed, it is well-known that the risk of unemployment is higher for persons with low incomes and it is unknown how much of this relationship is captured by the model. There are, however, a number of factors that increase our confidence that the benefit history parameters represent a causal effect of unemployment on the wage rate. Firstly, the estimations are done separately for persons with different education levels, which means, that – in terms of education – like are grouped with like and because of that, a large proportion of the correlation between unemployment risk and wage rates is accounted for. Also the inclusion of an extensive number of background variables serves to reduce the problem. Secondly, at least part of the potential endogeneity of the benefit variables is accounted for by the individual random effect (that is, the problem is that low wages and unemployment risk tend to coincide). More generally, the model’s dynamic parameters contribute to ensure that the measured effect of unemployment spells is adjusted for any simultaneity effects. In analysis not reported here, we test the endogeneity assumption by including lead variables for ‘benefit future’ that show a much weaker, albeit non-zero, relationship with wage rates than benefit history, which suggest that endogeneity is present but not a serious problem.
reduces the hourly wage rate by 5 per cent. Looking at the parameter estimates for the different groups the following can be summarised:

- The effect of periods on unemployment benefits and social security is negative for almost all the groups and tends to be stronger for persons with a tertiary education, particularly for social security recipients, (figure 2a, 2b and 2c).
- The effect of unemployment (figure 2a and 2b) on the wage rates is somewhat larger for males than for females.
- The wage rates for persons with a tertiary education are generally affected more than persons with basic schooling or a vocational education. Females with a medium length tertiary education are an exception – presumably because this group has many public sector employees (health and education).
- The impact of subsidised employment on the wage rates is generally negative and stronger for persons with a tertiary education. However, the effect is somewhat weaker than for inactive recipients of unemployment benefits and social security – again, females with a medium tertiary education stand out with a zero effect (figure 2d).
- The number of weeks on sickness benefits has a strong negative impact on the subsequent wage rates, and for males the effect is strongly related to education level (figure 2e).
- Periods on maternity leave benefits has little or no effect on subsequent wage rates (figure 2f).
- The labour market leave programmes of the nineties did not have a huge impact on subsequent wages. For persons with a long tertiary education there was even a small positive effect, presumably because many participated in educational programmes (figure 2g).

These results will of course have implications for the incomes of individuals who experience periods of non-employment. However, the economic cost in terms of lost productivity and forgone taxes etc. would also be substantial.

---

**Figure 2a**

Parameter estimates, share of weeks on unemployment benefits previous 3 years

<table>
<thead>
<tr>
<th></th>
<th>Primary</th>
<th>Secondary</th>
<th>Vocational</th>
<th>Short</th>
<th>Medium</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>-0.25</td>
<td>-0.2</td>
<td>-0.15</td>
<td>-0.1</td>
<td>-0.05</td>
<td>0</td>
</tr>
<tr>
<td>Females</td>
<td>-0.35</td>
<td>-0.3</td>
<td>-0.25</td>
<td>-0.2</td>
<td>-0.15</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

**Figure 2b**

Parameter estimates, share of weeks on social security (unemployed) previous 3 years

<table>
<thead>
<tr>
<th></th>
<th>Primary</th>
<th>Secondary</th>
<th>Vocational</th>
<th>Short</th>
<th>Medium</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>-0.25</td>
<td>-0.2</td>
<td>-0.15</td>
<td>-0.1</td>
<td>-0.05</td>
<td>0</td>
</tr>
<tr>
<td>Females</td>
<td>-0.35</td>
<td>-0.3</td>
<td>-0.25</td>
<td>-0.2</td>
<td>-0.15</td>
<td>-0.1</td>
</tr>
</tbody>
</table>
Figure 2c
Parameter estimates, share of weeks on social security (not unemployed) previous 3 years

Figure 2d
Parameter estimates, share of weeks in subsidised employment previous 3 years

Figure 2e
Parameter estimates, share of weeks on sickness benefits, previous 3 years

Figure 2f
Parameter estimates, share of weeks on maternity benefits, previous 3 years
Wage rate dynamics

The model’s comprehensive dynamic specification provides interesting insight into the dynamics of the wage process.

The $\gamma$ and the $\rho_2$ parameters represent distinctly different aspects of the wage dynamics. $\gamma$ represents the inter-temporal dependency of the wage rates, which is neither explained by the observed nor the unobserved factors (through the error process). $\rho_2$ represents inter-temporal correlation of the unobserved factors. The results show that $\gamma$ is generally lower than $\rho_2$ for both males and females (figure 3a and 3b). For males, $\gamma$ is generally increasing with the education level while $\rho_2$ tends to be lower for tertiary educations. There is no clear education pattern for females, but is a tendency to a trade-off between the $\gamma$ and the $\rho_2$ parameters.

The higher $\rho_2$-values imply that the dynamics of wage rates to a large extent are determined by inter-temporal constancy of unobserved factors. In contrast, a large $\gamma$ -value (males with a long tertiary education) imply a higher degree of year-to-year constancy, which is determined neither by observable nor unobservable wage determinants.

---

7 The $\rho_2$ parameter and the random effects together capture the dynamics of the unobserved factors embedded in the error term. The difference is that the random effect captures a permanent effect while the $\rho_2$ parameter capture a transient or year-to-year effects (see figure 5a).
The $\phi_1$ parameter is a permanent effect of the wage rate in a person’s first job on the wage rate in subsequent years. All groups exhibit this form of wage persistency, but the effect is clearly stronger for persons with a tertiary education while it is much less pronounced for persons with a primary or a vocational education (figure 4a). The implications are profound. An $\phi_1$-parameter in the range 0.3-0.4 implies that between 30 and 40 per cent of the unexplained part of the wage rate (below or above average) in the first year after leaving school or graduating tends to stick with a person in future years.

The importance of this effect can be illustrated by looking at the standard deviation of the first year error term, $\sigma_1$, which shows some variation across the groups. It is generally higher for males (around 0.19-0.23) than females (around 0.16-0.22) and slightly increasing with education level (figure 5a). As a rough guide, a starting wage rate above the average (conditional on the background variable) by one standard deviation of 0.2 and a $\phi_1$-parameter of 0.3 means that the wage rate will tend to be persistently higher by around 6 per cent over and above persons with similar characteristics – due to the starting wage alone.

---

Notice that $\phi_1$ only applies to persons who enter the panel when they finish an education – $\phi_2$ is the equivalent first-year effect for persons that enter the panel for other reasons.
The error term can be interpreted as a wage rate risk or an uncertainty about future wages faced by individuals (that is, as observed by the analyst). The individuals themselves may of course know more (and perhaps sometimes less) than the information available to the analyst in term of data and techniques. Nevertheless, the error term measures the accuracy with which the wage rate can be predicted from year to year and, with accumulated uncertainty, further out in the future. As such, the differences in error term variation across groups have important implications for risks and uncertainties for males versus females as well as for individuals with different educations.

The first year error term standard deviation is larger for males than for females for all education groups. Indeed, the female fraction of the total variation ranges from 24 per cent for medium length tertiary
educations to 40 per cent for short tertiary educations. This is not surprising, as the variation in wage rates is generally much larger for male. The message here is that observable factors cannot (fully) explain these gender differences. In other words, males may have higher wages than females, but they also face more uncertainty.

Figure 5a also exhibits substantial differences across education groups, albeit the patterns are less obvious than the gender differences. Females with a vocational or a medium length tertiary education stand out with very low error term variation – presumably due to the high rates of public sector employment in these groups. Conversely, males with a long tertiary education stand out with the largest error term variation.

The error term standard deviations for subsequent years also vary considerably across gender and education groups (figure 6b). Males consistently have higher variation than females and increases slightly with education level although, again, the females with a medium length tertiary education stand out at the low end.

The individual random effect represents another aspect of wage rate persistency. It is also sometimes referred to as a permanent luck effect although in the case of wage rates, this is arguably a somewhat misleading label unless one is willing to accept unobserved characteristics, such as effort and commitment as an outcome of ‘luck’ as much as cognitive abilities and intelligence.

The importance of the random effect as measured by its standard deviation \( \sigma^2 \) exhibits some variation across the gender and educational groups. Males generally have larger effects than females for all educational, except persons with long tertiary education who have almost equal size random effects for males and females (figure 6a).

The random effect should also be seen in the context of its fraction of the total unexplained variation (shown as a line in figure 6a). This fraction is between 50 and 60 per cent for all groups and a little higher for males than for females. As a consequence, a slightly larger fraction of the variation for males can be attributed to unobserved individual characteristics.
The first year error term standard deviation is much larger than the standard deviation for the subsequent years, cf. figure 5a, 5b and 6b. This is a natural consequence of the model structure having a first year wage rate determined only by the background variables (1). In contrast, the wage rate in subsequent years is determined with higher accuracy because it depends on the outcome in previous years (2). The first year salary is particularly important because it works through to subsequent years both through a fading effect, the $\gamma$-parameter, and a permanent effect, the $\phi$-parameter (notice that $\rho_1$ is close to zero for all groups).

The Gaussian error term and random effect distributions

In spite of the wage rate model’s extensive explanatory power, the unexplained wage rate variation imbedded in the error terms and the random effects ((6), (7) and (8)) still represent a large proportion of the overall dispersion of the wage rates. Hence, it is important that the distributions of these terms are modelled in a realistic manner. In SMILE this is done by applying mixed Gaussian distributions – the estimated parameters for these distributions are shown in table A3 (Appendix 1).

Mixed Gaussian distributions have the ability to emulate a wide range of distribution shapes. The following figures illustrate this for males with a primary school education. The figures compare the actual distributions of the model’s the random effect (figure 7a) and its three error terms (figure 7b, 7c and 7d) with the modelled distribution using either a single normal distribution of a mixture of several normal distributions.

Comparing the residuals’ actual distributions with normal distributions with identical standard deviations shows that the actual distributions are strongly leptokurtic and in three cases they are also visibly skewed. By implication, the residuals are not normally distributed and applying a simple normal distribution for the simulation of wage rates would lead to a serious bias in the distribution of simulated wage rates. In contrast, the distributions obtained by applying a mixture of normal distributions are so close to the actual distributions that only a small sample size tells them apart.
The consequences of applying a single normal distribution to the simulation of wage rates are unpredictable, but surely the distribution of incomes would be unrealistic. The bias for the random effect (figure 7a) and the year 2+ error term (figure 7b) would push the earned incomes of a large number of persons away from the area around the middle of the distribution and, at the same time, reducing the number of persons with more extreme outcomes.

Failing to account for the skewed distribution of the random effect (figure 7a) and the first year error term (figure 7c and 7d) would push mass of incomes to the right and thereby cause a misleading picture of the overall income distribution.

Other background variables

A number of the results for the background variables have important interpretations that can be summarised by the following points:

- Cohabitation status affect male wages but not female wages: single males have 3 to 6 per cent lower 1st year wage rate and 1 to 2.5 per cent lower wage rate in subsequent years than cohabiting males.
- Children on affect wage rates for males with primary school education and only through 1st year effect. However, the effect is rather large and suggests that males with only primary school education have wages that start roughly 19 per cent below their fertile counterparts.
- Birthplace matters for wage rates. For most groups wages rates tend to be higher for persons from other western countries tend to have higher wage rates than Danish-born persons (by around 2 to 7 per cent) and persons from non-western countries tend to have lower wage rates than Danish-born persons.
- The parents’ education has a small influence on a person’s wage rate. The somewhat weak direct link between parents’ education and wage rates is perhaps surprising, but the effect is conditional on a person’s own education, which is accounted for by grouping. Indeed, we know from other research that a person’s own educational choice is influenced by the parents’ education – especially the mother’s education. The important finding here is that parents’ education may affect
their children’s income through their choice of education, but there is only a small effect on postgraduate wages beyond educational choice.

- Secondary school grades have a substantial positive impact on the wage rates for all education groups (except of course primary school) by between 3.5 and 5.5 per cent per point.
References


Baekgaard, H. (2014b) A Dynamic Model for Worked Hours for Employees, DREAM report (forthcoming, will be available from www.dreammodel.dk/SMILE)


Appendix 1 Estimation results

The following tables show the parameter results of the estimation of the model (1) - (5). Many background variables are self-explanatory. Nevertheless, the detailed descriptions below provide more precise definitions as a reference guide.

Dummy_1995 - Dummy_2010: year-dummies

Age: a person's age 1\textsuperscript{st} January

Age squared: ‘Age’ squared divided by 100

Children present: 1, if there are children aged 0-17 in the family

Single: 1, if a person lives in a one-adult family

COB west: 1, if the person was born in a western country other than Denmark

COB non-west: 1, if the person was born in a non-western country

COB Nordic: 1, if the person was born in a Nordic country other than Denmark

COB w Euro: 1, if the person was born in a west European country other than the Nordic countries

COB other Euro: 1, if the person was born in a other European countries

COB Africa: 1, if the person was born in Africa

COB Lat. America: 1, if the person was born in Latin America

COB e. Asia den.: 1, if the person was born in east Asia (more developed)

COB e. Asia n den.: 1, if the person was born in east Asia (less developed)

COB w. Asia.: 1, if the person was born in west Asia (less developed)

YSA non-west: the number of years since arrival (persons born in a non-western country)

YSA west: the number of years since arrival (persons born in a western country)

Desc. west: 1, if the person is descendant from a western country

Desc. non-west: 1, if the person is descendant from a non-western country
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Att. Ed. primary</td>
<td>1, if the person is attending primary school</td>
</tr>
<tr>
<td>Att. Ed. secondary</td>
<td>1, if the person is attending secondary school</td>
</tr>
<tr>
<td>Att. Ed. vocational</td>
<td>1, if the person is attending a vocational education</td>
</tr>
<tr>
<td>Att. Ed. tert. short</td>
<td>1, if the person is attending a short tertiary education</td>
</tr>
<tr>
<td>Att. Ed. tert. Med.</td>
<td>1, if the person is attending a medium tertiary education</td>
</tr>
<tr>
<td>Att. Ed. tert. long</td>
<td>1, if the person is attending a long tertiary education</td>
</tr>
<tr>
<td>Dad's educ 1</td>
<td>1, if the father’s highest education is primary school</td>
</tr>
<tr>
<td>Dad's educ 2</td>
<td>1, if the father’s highest education is secondary school</td>
</tr>
<tr>
<td>Dad's educ 3</td>
<td>1, if the father’s highest education is vocational</td>
</tr>
<tr>
<td>Dad's educ 4</td>
<td>1, if the father’s highest education is short tertiary</td>
</tr>
<tr>
<td>Dad's educ 5</td>
<td>1, if the father’s highest education is medium tertiary</td>
</tr>
<tr>
<td>Dad's educ 6</td>
<td>1, if the father’s highest education is long tertiary</td>
</tr>
<tr>
<td>Mum's educ 1</td>
<td>1, if the mother’s highest education is primary school</td>
</tr>
<tr>
<td>Mum's educ 2</td>
<td>1, if the mother’s highest education is secondary school</td>
</tr>
<tr>
<td>Mum's educ 3</td>
<td>1, if the mother’s highest education is vocational</td>
</tr>
<tr>
<td>Mum's educ 4</td>
<td>1, if the mother’s highest education is short tertiary</td>
</tr>
<tr>
<td>Mum's educ 5</td>
<td>1, if the mother’s highest education is medium tertiary</td>
</tr>
<tr>
<td>Mum's educ 6</td>
<td>1, if the mother’s highest education is long tertiary</td>
</tr>
<tr>
<td>Experience</td>
<td>number of years employment experience</td>
</tr>
<tr>
<td>Experience, square root</td>
<td>square root of 'Experience'</td>
</tr>
<tr>
<td>Graduate</td>
<td>1, if first year in employment after completing highest education</td>
</tr>
<tr>
<td>First job</td>
<td>1, if first employment after completing highest education</td>
</tr>
<tr>
<td>Secondary grade</td>
<td>average grade from secondary school</td>
</tr>
</tbody>
</table>
Secondary grade NA: 1, if the person completed secondary school but has no grade is available

1, UE Benefits: =1, if having received unemployment benefits in the current year

1, Soc. Sec. UE: =1, if having received social security as unemployed in the current year

1, Soc. Sec. nUE: =1, if having received social security as not unemployed in the current year

1, subs. empl.: =1, if having worked in a job with salary subsidy in the current year

1, sickness benefits: =1, if having received sickness benefits in the current year

1, maternity leave: =1, if having received maternity leave benefits in the current year

1, other leave: =1, if having received other leave benefits in the current year

1, not in lab. force: =1, if having been out of the labour force without benefits in the current year

Wks UE Benefits: the share of the number of weeks if having received unemployment benefits in the previous three years

Wks S. Sec. UE the share of the number of weeks if having received social security as unemployed in the previous three years

Wks S. Sec. nUE the share of the number of weeks if having received social security as not unemployed in the previous three years

Wks subs. empl.: the share of the number of weeks worked in a job with salary subsidy in the previous three years

Wks disab. ben.: the share of the number of weeks having received disability benefits in the previous three years

Wks mater. leave.: the share of the number of weeks having received maternity leave benefits in the previous three years

Wks other leave.: the share of the number of weeks having received other leave benefits in the previous three years

Wks not in lab f.: the share of the number of weeks having been out of the labour force without benefits in the previous three years
## Table A1

**Estimation results for gender/education groups – first year**

<table>
<thead>
<tr>
<th></th>
<th>Primary school</th>
<th>Secondary school</th>
<th>Vocational</th>
<th>Short tertiary</th>
<th>Medium tertiary</th>
<th>Long tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Dummy_1995</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dummy_1996</td>
<td>-0.012</td>
<td>-0.009</td>
<td>-</td>
<td>-0.014</td>
<td>-</td>
<td>-0.007</td>
</tr>
<tr>
<td>Dummy_1997</td>
<td>-0.029</td>
<td>-0.031</td>
<td>-</td>
<td>-0.032</td>
<td>-</td>
<td>-0.042</td>
</tr>
<tr>
<td>Dummy_1998</td>
<td>-0.035</td>
<td>-0.042</td>
<td>-</td>
<td>-0.043</td>
<td>-</td>
<td>-0.037</td>
</tr>
<tr>
<td>Dummy_1999</td>
<td>-0.043</td>
<td>-0.048</td>
<td>-</td>
<td>-0.045</td>
<td>-</td>
<td>-0.048</td>
</tr>
<tr>
<td>Dummy_2000</td>
<td>-0.036</td>
<td>-0.056</td>
<td>-</td>
<td>-0.072</td>
<td>-</td>
<td>-0.043</td>
</tr>
<tr>
<td>Dummy_2001</td>
<td>-0.034</td>
<td>-0.052</td>
<td>-</td>
<td>-0.095</td>
<td>-</td>
<td>-0.058</td>
</tr>
<tr>
<td>Dummy_2002</td>
<td>-0.069</td>
<td>-0.065</td>
<td>-</td>
<td>-0.113</td>
<td>-</td>
<td>-0.070</td>
</tr>
<tr>
<td>Dummy_2003</td>
<td>-0.063</td>
<td>-0.058</td>
<td>-</td>
<td>-0.102</td>
<td>-</td>
<td>-0.052</td>
</tr>
<tr>
<td>Dummy_2004</td>
<td>-0.065</td>
<td>-0.074</td>
<td>-</td>
<td>-0.095</td>
<td>-</td>
<td>-0.069</td>
</tr>
<tr>
<td>Dummy_2005</td>
<td>-0.056</td>
<td>-0.085</td>
<td>-</td>
<td>-0.070</td>
<td>-</td>
<td>-0.051</td>
</tr>
<tr>
<td>Dummy_2006</td>
<td>-0.132</td>
<td>-0.105</td>
<td>-</td>
<td>-0.032</td>
<td><strong>-0.041</strong></td>
<td>-0.127</td>
</tr>
<tr>
<td>Dummy_2007</td>
<td>-0.146</td>
<td>-0.119</td>
<td><strong>-0.002</strong></td>
<td><strong>-0.057</strong></td>
<td>-0.117</td>
<td>-0.090</td>
</tr>
<tr>
<td>Dummy_2008</td>
<td>-0.156</td>
<td>-0.118</td>
<td>0.047</td>
<td>0.032</td>
<td><strong>-0.07</strong></td>
<td>-0.13</td>
</tr>
<tr>
<td>Dummy_2009</td>
<td>-0.200</td>
<td>-0.162</td>
<td>0.049</td>
<td>0.034</td>
<td><strong>-0.06</strong></td>
<td>-0.137</td>
</tr>
<tr>
<td>Dummy_2010</td>
<td>-0.212</td>
<td>-0.161</td>
<td>0.042</td>
<td>0.022</td>
<td>-0.083</td>
<td>-0.162</td>
</tr>
<tr>
<td>Age</td>
<td>0.019</td>
<td>0.012</td>
<td>0.052</td>
<td>0.038</td>
<td>0.022</td>
<td>0.015</td>
</tr>
<tr>
<td>Age squared</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Children present</td>
<td>0.191</td>
<td><strong>0.013</strong></td>
<td><strong>0.095</strong></td>
<td><strong>0.063</strong></td>
<td><strong>-0.035</strong></td>
<td><strong>-0.023</strong></td>
</tr>
<tr>
<td>Single</td>
<td>-0.040</td>
<td>0.008</td>
<td>-0.058</td>
<td>0.009</td>
<td>-0.034</td>
<td><strong>0.005</strong></td>
</tr>
<tr>
<td>COB west</td>
<td>-</td>
<td>-</td>
<td><strong>0.003</strong></td>
<td>-0.045</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COB non-west</td>
<td>-</td>
<td>-</td>
<td>-0.054</td>
<td>-0.037</td>
<td>-0.091</td>
<td>-0.06</td>
</tr>
<tr>
<td>COB Nordic</td>
<td>-</td>
<td>0.041</td>
<td>0.086</td>
<td>-0.063</td>
<td>-</td>
<td><strong>0.018</strong></td>
</tr>
<tr>
<td>COB w. Euro</td>
<td>-</td>
<td>0.036</td>
<td>0.043</td>
<td>0.071</td>
<td>-</td>
<td>0.055</td>
</tr>
<tr>
<td>COB other Euro</td>
<td>0.004</td>
<td>0.009</td>
<td>0.017</td>
<td>0.013</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td>COB Africa</td>
<td>-0.089</td>
<td>-0.280</td>
<td>-0.095</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB Lat. America</td>
<td>0.092</td>
<td>-0.070</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB e. Asia dev.</td>
<td>**0.083</td>
<td>-0.021</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB w. Asia</td>
<td>0.059</td>
<td>-0.243</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB NA</td>
<td>-0.004</td>
<td>0.002</td>
<td>**0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YSA non-west</td>
<td>**0.024</td>
<td>0.011</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desc. west</td>
<td>-0.015</td>
<td>-0.012</td>
<td>-0.091</td>
<td>0.018</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Att. ed. primary</td>
<td>-0.040</td>
<td>-0.057</td>
<td>**0.016</td>
<td>-0.052</td>
<td>-0.012</td>
<td>-0.091</td>
</tr>
<tr>
<td>Att. ed. Secondary</td>
<td>-0.076</td>
<td>-0.094</td>
<td>**0.044</td>
<td>-0.211</td>
<td>**0.079</td>
<td>- **0.010</td>
</tr>
<tr>
<td>Att. ed. tert. short</td>
<td>-0.087</td>
<td>-0.046</td>
<td>**0.068</td>
<td>**0.028</td>
<td>-0.124</td>
<td>-0.162</td>
</tr>
<tr>
<td>Att. ed. tert. med.</td>
<td>-0.097</td>
<td>-0.046</td>
<td>**0.055</td>
<td>-0.119</td>
<td>-0.139</td>
<td>-0.114</td>
</tr>
<tr>
<td>Att. ed. tert. long</td>
<td>**0.036</td>
<td>-0.113</td>
<td>-0.046</td>
<td>**0.024</td>
<td>-0.087</td>
<td>**0.035</td>
</tr>
<tr>
<td>Dad’s educ. NA</td>
<td>**0.013</td>
<td>**0.009</td>
<td>**0.010</td>
<td>-0.015</td>
<td>-0.18</td>
<td>-0.015</td>
</tr>
<tr>
<td>Dad’s educ. 1</td>
<td>0.013</td>
<td>0.014</td>
<td>0.017</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dad’s educ. 2</td>
<td>0.013</td>
<td>0.014</td>
<td>0.017</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dad’s educ. 3</td>
<td>0.013</td>
<td>0.014</td>
<td>0.021</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dad’s educ. 4</td>
<td>0.013</td>
<td>0.014</td>
<td>0.021</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dad’s educ. 5</td>
<td>0.013</td>
<td>0.014</td>
<td>0.017</td>
<td>**0.006</td>
<td>0.023</td>
<td>0.011</td>
</tr>
<tr>
<td>Dad’s educ. 6</td>
<td>0.013</td>
<td>0.014</td>
<td>0.017</td>
<td>**0.006</td>
<td>0.023</td>
<td>0.011</td>
</tr>
<tr>
<td>Mum’s educ. NA</td>
<td>-0.022</td>
<td>-0.029</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mum’s educ. 1</td>
<td>-0.026</td>
<td>-0.017</td>
<td>-0.019</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mum’s educ. 2</td>
<td>0.015</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mum’s educ. 3</td>
<td>0.015</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mum’s educ. 4</td>
<td>0.015</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mum’s educ. 5</td>
<td>0.015</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mum’s educ. 6</td>
<td>0.015</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Experience</td>
<td>**0.060</td>
<td>**0.022</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Exp. square root</td>
<td>0.017</td>
<td>-0.014</td>
<td>0.014</td>
<td>**0.01</td>
<td>0.056</td>
<td>**0.066</td>
</tr>
<tr>
<td>Graduate</td>
<td>**0.047</td>
<td>-0.035</td>
<td>**0.013</td>
<td>-0.017</td>
<td>**0.016</td>
<td>**0.007</td>
</tr>
</tbody>
</table>
Note: "**" indicates that the estimate is not significant at a 5% level.

<table>
<thead>
<tr>
<th></th>
<th>-</th>
<th>-</th>
<th>0.105</th>
<th>0.176</th>
<th>**0.046</th>
<th>0.173</th>
<th>0.129</th>
<th>0.237</th>
<th>0.134</th>
<th>0.120</th>
<th>0.205</th>
<th>0.273</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second. grade NA</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary grade</td>
<td>-</td>
<td>-</td>
<td>0.015</td>
<td>0.027</td>
<td>**0.015</td>
<td>0.030</td>
<td>0.02</td>
<td>0.035</td>
<td>0.020</td>
<td>0.020</td>
<td>0.023</td>
<td>0.031</td>
</tr>
<tr>
<td>1, UE benefits</td>
<td>-0.033</td>
<td>-0.035</td>
<td>-0.073</td>
<td>-0.035</td>
<td>-0.037</td>
<td>-0.046</td>
<td>-0.092</td>
<td>-0.074</td>
<td>-0.055</td>
<td>-0.020</td>
<td>-0.121</td>
<td>-0.089</td>
</tr>
<tr>
<td>1, Soc. Sec. UE</td>
<td>-0.019</td>
<td>-0.012</td>
<td>-0.024</td>
<td>-0.012</td>
<td>**-0.004</td>
<td>-0.032</td>
<td>-0.061</td>
<td>-0.043</td>
<td>-0.072</td>
<td>**-0.003</td>
<td>-0.120</td>
<td>-0.080</td>
</tr>
<tr>
<td>1, Soc. Sec. nUE</td>
<td>-0.027</td>
<td>-0.023</td>
<td>-0.062</td>
<td>-0.021</td>
<td>-0.062</td>
<td>-0.037</td>
<td>-0.067</td>
<td>-0.048</td>
<td>-0.029</td>
<td>-</td>
<td>-0.045</td>
<td>-0.038</td>
</tr>
<tr>
<td>1, subs. empl.</td>
<td>-0.076</td>
<td>-0.047</td>
<td>-0.052</td>
<td>-0.045</td>
<td>-0.039</td>
<td>-0.024</td>
<td>-</td>
<td>-</td>
<td>**0.006</td>
<td>-</td>
<td>**0.009</td>
<td>**-0.033</td>
</tr>
<tr>
<td>1, sickness benefits</td>
<td>-0.015</td>
<td>-</td>
<td>-0.045</td>
<td>-0.016</td>
<td>**-0.003</td>
<td>-</td>
<td>-0.056</td>
<td>-0.025</td>
<td>-0.059</td>
<td>**-0.005</td>
<td>-0.049</td>
<td>-0.011</td>
</tr>
<tr>
<td>1, maternity leave</td>
<td>**0.023</td>
<td>0.032</td>
<td>**0.025</td>
<td>0.024</td>
<td>-</td>
<td>**0.009</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>**-0.012</td>
<td>**-0.004</td>
<td>-0.020</td>
</tr>
<tr>
<td>1, other leave</td>
<td>-0.049</td>
<td>**-0.011</td>
<td>-0.056</td>
<td>-0.041</td>
<td>-0.047</td>
<td>-0.030</td>
<td>-0.028</td>
<td>-0.034</td>
<td>-0.144</td>
<td>-0.036</td>
<td>-0.103</td>
<td>-0.086</td>
</tr>
<tr>
<td>1, not in lab. force</td>
<td>**-0.023</td>
<td>**-0.005</td>
<td>-0.113</td>
<td>-</td>
<td>-0.056</td>
<td>**-0.009</td>
<td>**-0.045</td>
<td>**-0.035</td>
<td>**-0.075</td>
<td>**-0.032</td>
<td>**-0.068</td>
<td>-0.137</td>
</tr>
<tr>
<td>Wks, UE benefits</td>
<td>-0.123</td>
<td>-0.089</td>
<td>-0.218</td>
<td>-0.144</td>
<td>-0.194</td>
<td>-0.118</td>
<td>-0.182</td>
<td>-0.134</td>
<td>-0.240</td>
<td>-0.101</td>
<td>-0.346</td>
<td>-0.219</td>
</tr>
<tr>
<td>Wks, S. sec. UE</td>
<td>-0.064</td>
<td>-0.069</td>
<td>-0.234</td>
<td>-0.207</td>
<td>-0.231</td>
<td>-0.068</td>
<td>-0.275</td>
<td>-0.202</td>
<td>-0.322</td>
<td>-0.233</td>
<td>-0.459</td>
<td>-0.453</td>
</tr>
<tr>
<td>Wks, S. Sec. nUE</td>
<td>-0.061</td>
<td>-0.036</td>
<td>-0.238</td>
<td>-0.144</td>
<td>-0.160</td>
<td>-0.047</td>
<td>-0.218</td>
<td>-0.133</td>
<td>-0.402</td>
<td>-0.222</td>
<td>-0.587</td>
<td>-0.399</td>
</tr>
<tr>
<td>Wks, subs. empl.</td>
<td>**-0.048</td>
<td>**0.025</td>
<td>-0.172</td>
<td>-</td>
<td>**-0.010</td>
<td>-</td>
<td>-0.068</td>
<td>-0.081</td>
<td>-0.128</td>
<td>-0.053</td>
<td>-0.199</td>
<td>**-0.044</td>
</tr>
<tr>
<td>Wks, sickness. ben.</td>
<td>**-0.064</td>
<td>**-0.033</td>
<td>-0.329</td>
<td>-0.170</td>
<td>-0.164</td>
<td>-0.081</td>
<td>-0.348</td>
<td>-0.169</td>
<td>-0.542</td>
<td>-0.149</td>
<td>-0.472</td>
<td>**-0.062</td>
</tr>
<tr>
<td>Wks, mater. leave</td>
<td>**0.393</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.103</td>
<td>-</td>
<td>**-0.013</td>
<td>**0.210</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wks, other leave</td>
<td>-0.136</td>
<td>-0.044</td>
<td>-0.284</td>
<td>-0.152</td>
<td>-0.255</td>
<td>-0.105</td>
<td>-0.136</td>
<td>-0.125</td>
<td>-0.306</td>
<td>-0.145</td>
<td>-0.179</td>
<td>-0.127</td>
</tr>
<tr>
<td>Wks, not in lab. f.</td>
<td>-0.131</td>
<td>-0.092</td>
<td>-0.162</td>
<td>-0.248</td>
<td>-0.202</td>
<td>-0.086</td>
<td>-0.266</td>
<td>**-0.093</td>
<td>-0.172</td>
<td>-0.104</td>
<td>-0.280</td>
<td>-</td>
</tr>
</tbody>
</table>
Table A2

Estimation results for gender/education groups – dynamic parameters

<table>
<thead>
<tr>
<th></th>
<th>Primary school</th>
<th>Secondary school</th>
<th>Vocational</th>
<th>Short tertiary</th>
<th>Medium tertiary</th>
<th>Long tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.109</td>
<td>0.123</td>
<td>0.059</td>
<td>0.193</td>
<td>0.173</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.004)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>$\phi_1$ (Graduates)</td>
<td>0.122</td>
<td>0.123</td>
<td>0.310</td>
<td>0.186</td>
<td>0.138</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.025)</td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.014)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>$\phi_2$ (Others)</td>
<td>0.400</td>
<td>0.331</td>
<td>0.496</td>
<td>0.355</td>
<td>0.418</td>
<td>0.386</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.014</td>
<td>0.003</td>
<td>-0.003</td>
<td>-0.017</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.333</td>
<td>0.326</td>
<td>0.451</td>
<td>0.279</td>
<td>0.296</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>0.036</td>
<td>0.028</td>
<td>0.035</td>
<td>0.025</td>
<td>0.053</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$\sigma_{11}$ (Graduates)</td>
<td>0.057</td>
<td>0.036</td>
<td>0.078</td>
<td>0.045</td>
<td>0.071</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$\sigma_{12}$ (Others)</td>
<td>0.014</td>
<td>0.009</td>
<td>0.021</td>
<td>0.015</td>
<td>0.015</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.020</td>
<td>0.012</td>
<td>0.032</td>
<td>0.015</td>
<td>0.020</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>
### Table A3

<table>
<thead>
<tr>
<th></th>
<th>Primary school</th>
<th>Secondary school</th>
<th>Vocational</th>
<th>Short tertiary</th>
<th>Medium tertiary</th>
<th>Long tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>$\epsilon_{11}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{11}$</td>
<td>-0.079</td>
<td>-0.077</td>
<td>-0.143</td>
<td>-0.111</td>
<td>-0.317</td>
<td>-0.192</td>
</tr>
<tr>
<td>$\alpha_{12}$</td>
<td>0.101</td>
<td>0.046</td>
<td>-0.023</td>
<td>-0.002</td>
<td>0.010</td>
<td>0.012</td>
</tr>
<tr>
<td>$\alpha_{13}$</td>
<td>0.568</td>
<td>0.429</td>
<td>0.009</td>
<td>-0.004</td>
<td>0.366</td>
<td>0.065</td>
</tr>
<tr>
<td>$\alpha_{14}$</td>
<td>0.212</td>
<td>0.287</td>
<td>0.378</td>
<td>0.210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{11}$</td>
<td>0.010</td>
<td>0.006</td>
<td>0.004</td>
<td>0.003</td>
<td>0.006</td>
<td>0.010</td>
</tr>
<tr>
<td>$\sigma_{12}$</td>
<td>0.023</td>
<td>0.014</td>
<td>0.003</td>
<td>0.002</td>
<td>0.038</td>
<td>0.006</td>
</tr>
<tr>
<td>$\sigma_{13}$</td>
<td>0.089</td>
<td>0.078</td>
<td>0.018</td>
<td>0.015</td>
<td>0.100</td>
<td>0.021</td>
</tr>
<tr>
<td>$\sigma_{14}$</td>
<td>0.077</td>
<td>0.069</td>
<td>0.118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_{11}$</td>
<td>0.641</td>
<td>0.532</td>
<td>0.251</td>
<td>0.240</td>
<td>0.080</td>
<td>0.207</td>
</tr>
<tr>
<td>$p_{12}$</td>
<td>0.324</td>
<td>0.408</td>
<td>0.231</td>
<td>0.253</td>
<td>0.871</td>
<td>0.340</td>
</tr>
<tr>
<td>$p_{13}$</td>
<td>0.035</td>
<td>0.061</td>
<td>0.340</td>
<td>0.407</td>
<td>0.049</td>
<td>0.418</td>
</tr>
<tr>
<td>$p_{14}$</td>
<td>0.177</td>
<td>0.100</td>
<td>0.034</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\epsilon_{21}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{21}$</td>
<td>-0.104</td>
<td>-0.126</td>
<td>-0.095</td>
<td>-0.098</td>
<td>-0.115</td>
<td>-0.157</td>
</tr>
<tr>
<td>$\alpha_{22}$</td>
<td>0.011</td>
<td>-0.017</td>
<td>-0.018</td>
<td>-0.045</td>
<td>0.025</td>
<td>-0.021</td>
</tr>
<tr>
<td>$\alpha_{23}$</td>
<td>0.385</td>
<td>0.057</td>
<td>0.330</td>
<td>0.077</td>
<td>0.408</td>
<td>0.075</td>
</tr>
<tr>
<td>$\alpha_{24}$</td>
<td>0.409</td>
<td></td>
<td>0.421</td>
<td>0.371</td>
<td>0.320</td>
<td>0.243</td>
</tr>
<tr>
<td>$\sigma_{21}$</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.017</td>
<td>0.025</td>
<td>0.018</td>
</tr>
<tr>
<td>$\sigma_{22}$</td>
<td>0.044</td>
<td>0.007</td>
<td>0.009</td>
<td>0.009</td>
<td>0.060</td>
<td>0.009</td>
</tr>
<tr>
<td>$\sigma_{23}$</td>
<td>0.136</td>
<td>0.031</td>
<td>0.161</td>
<td>0.058</td>
<td>0.147</td>
<td>0.029</td>
</tr>
<tr>
<td>$\sigma_{24}$</td>
<td>0.120</td>
<td>0.167</td>
<td>0.117</td>
<td>0.127</td>
<td>0.144</td>
<td>0.113</td>
</tr>
<tr>
<td>$p_{21}$</td>
<td>0.369</td>
<td>0.274</td>
<td>0.324</td>
<td>0.254</td>
<td>0.397</td>
<td>0.258</td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>0.544</td>
<td>0.328</td>
<td>0.552</td>
<td>0.358</td>
<td>0.524</td>
<td>0.302</td>
</tr>
<tr>
<td>$p_{23}$</td>
<td>0.087</td>
<td>0.350</td>
<td>0.125</td>
<td>0.353</td>
<td>0.079</td>
<td>0.391</td>
</tr>
<tr>
<td>$p_{24}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### $\varepsilon_2$ (subsequent year error term)

<table>
<thead>
<tr>
<th></th>
<th>$p_{24}$</th>
<th>0.048</th>
<th>0.035</th>
<th>0.049</th>
<th>0.073</th>
<th>0.245</th>
<th>0.261</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{21}$</td>
<td>-0.006</td>
<td>-0.005</td>
<td>-0.017</td>
<td>-0.012</td>
<td>0.002</td>
<td>-0.008</td>
<td>-0.006</td>
</tr>
<tr>
<td>$\alpha_{22}$</td>
<td>-0.004</td>
<td>-0.025</td>
<td>-0.035</td>
<td>-0.020</td>
<td>-0.002</td>
<td>-0.023</td>
<td>-0.003</td>
</tr>
<tr>
<td>$\alpha_{23}$</td>
<td>0.045</td>
<td>0.029</td>
<td>0.052</td>
<td>0.038</td>
<td>0.014</td>
<td>0.028</td>
<td>0.008</td>
</tr>
<tr>
<td>$\alpha_{24}$</td>
<td>0.069</td>
<td>0.015</td>
<td>0.028</td>
<td>0.037</td>
<td>0.022</td>
<td>0.009</td>
<td>0.038</td>
</tr>
<tr>
<td>$\sigma_{21}$</td>
<td>0.021</td>
<td>0.014</td>
<td>0.034</td>
<td>0.021</td>
<td>0.021</td>
<td>0.015</td>
<td>0.081</td>
</tr>
<tr>
<td>$\sigma_{22}$</td>
<td>0.006</td>
<td>0.003</td>
<td>0.009</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
<td>0.005</td>
</tr>
<tr>
<td>$\sigma_{23}$</td>
<td>0.079</td>
<td>0.003</td>
<td>0.016</td>
<td>0.008</td>
<td>0.088</td>
<td>0.003</td>
<td>0.019</td>
</tr>
<tr>
<td>$\sigma_{24}$</td>
<td>0.087</td>
<td>0.095</td>
<td>0.088</td>
<td>0.077</td>
<td>0.080</td>
<td>0.080</td>
<td>0.080</td>
</tr>
<tr>
<td>$p_{21}$</td>
<td>0.520</td>
<td>0.499</td>
<td>0.329</td>
<td>0.461</td>
<td>0.478</td>
<td>0.470</td>
<td>0.110</td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>0.407</td>
<td>0.222</td>
<td>0.275</td>
<td>0.237</td>
<td>0.429</td>
<td>0.230</td>
<td>0.502</td>
</tr>
<tr>
<td>$p_{23}$</td>
<td>0.073</td>
<td>0.236</td>
<td>0.250</td>
<td>0.225</td>
<td>0.093</td>
<td>0.237</td>
<td>0.388</td>
</tr>
<tr>
<td>$p_{24}$</td>
<td>0.043</td>
<td>0.145</td>
<td>0.078</td>
<td>0.064</td>
<td>0.084</td>
<td>0.084</td>
<td>0.084</td>
</tr>
</tbody>
</table>

### $\tau$ (the random effect)

<table>
<thead>
<tr>
<th></th>
<th>$p_{24}$</th>
<th>0.014</th>
<th>0.005</th>
<th>0.010</th>
<th>-0.046</th>
<th>0.023</th>
<th>0.008</th>
<th>0.017</th>
<th>0.014</th>
<th>0.022</th>
<th>-0.012</th>
<th>-0.012</th>
<th>0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{21}$</td>
<td>-0.040</td>
<td>-0.020</td>
<td>-0.038</td>
<td>-0.015</td>
<td>-0.038</td>
<td>-0.027</td>
<td>-0.042</td>
<td>-0.032</td>
<td>-0.037</td>
<td>0.043</td>
<td>0.010</td>
<td>-0.046</td>
<td></td>
</tr>
<tr>
<td>$\alpha_{22}$</td>
<td>0.041</td>
<td>0.036</td>
<td>0.045</td>
<td>0.046</td>
<td>0.028</td>
<td>0.033</td>
<td>0.032</td>
<td>0.030</td>
<td>0.036</td>
<td>0.019</td>
<td>-0.003</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>$\alpha_{23}$</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{21}$</td>
<td>0.010</td>
<td>0.007</td>
<td>0.018</td>
<td>0.004</td>
<td>0.012</td>
<td>0.007</td>
<td>0.010</td>
<td>0.008</td>
<td>0.019</td>
<td>0.005</td>
<td>0.009</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{22}$</td>
<td>0.003</td>
<td>0.002</td>
<td>0.005</td>
<td>0.001</td>
<td>0.004</td>
<td>0.002</td>
<td>0.004</td>
<td>0.004</td>
<td>0.005</td>
<td>0.029</td>
<td>0.108</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{23}$</td>
<td>0.047</td>
<td>0.035</td>
<td>0.068</td>
<td>0.006</td>
<td>0.046</td>
<td>0.032</td>
<td>0.038</td>
<td>0.033</td>
<td>0.075</td>
<td>0.035</td>
<td>0.033</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{24}$</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td>$p_{21}$</td>
<td>0.471</td>
<td>0.503</td>
<td>0.486</td>
<td>0.364</td>
<td>0.441</td>
<td>0.496</td>
<td>0.446</td>
<td>0.452</td>
<td>0.510</td>
<td>0.775</td>
<td>0.729</td>
<td>0.184</td>
<td></td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>0.348</td>
<td>0.364</td>
<td>0.338</td>
<td>0.200</td>
<td>0.384</td>
<td>0.346</td>
<td>0.338</td>
<td>0.365</td>
<td>0.382</td>
<td>0.225</td>
<td>0.022</td>
<td>0.297</td>
<td></td>
</tr>
<tr>
<td>$p_{23}$</td>
<td>0.181</td>
<td>0.133</td>
<td>0.176</td>
<td>0.252</td>
<td>0.175</td>
<td>0.158</td>
<td>0.216</td>
<td>0.183</td>
<td>0.108</td>
<td>0.186</td>
<td>0.213</td>
<td>0.324</td>
<td></td>
</tr>
<tr>
<td>$p_{24}$</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td>0.184</td>
<td></td>
</tr>
</tbody>
</table>
Table A4

Estimation results for gender/education groups – subsequent years

<table>
<thead>
<tr>
<th></th>
<th>Primary school</th>
<th>Secondary school</th>
<th>Vocational</th>
<th>Short tertiary</th>
<th>Medium tertiary</th>
<th>Long tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Dummy_1997</td>
<td>0.008</td>
<td>0.007</td>
<td>-</td>
<td>0.006</td>
<td>0.015</td>
<td>0.007</td>
</tr>
<tr>
<td>Dummy_1998</td>
<td><strong>0.003</strong></td>
<td><strong>0.003</strong></td>
<td>-</td>
<td><strong>0.000</strong></td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>Dummy_1999</td>
<td>-0.009</td>
<td>-0.005</td>
<td>-0.013</td>
<td>-0.006</td>
<td><strong>-0.001</strong></td>
<td>-0.008</td>
</tr>
<tr>
<td>Dummy_2000</td>
<td>-0.008</td>
<td>-0.007</td>
<td>-0.017</td>
<td>-0.016</td>
<td><strong>0.002</strong></td>
<td>-0.009</td>
</tr>
<tr>
<td>Dummy_2001</td>
<td>-0.006</td>
<td>-0.010</td>
<td>-0.032</td>
<td>-0.018</td>
<td>-0.011</td>
<td>-0.016</td>
</tr>
<tr>
<td>Dummy_2002</td>
<td>-0.009</td>
<td>-0.016</td>
<td>-0.034</td>
<td>-0.021</td>
<td>-0.017</td>
<td>-0.004</td>
</tr>
<tr>
<td>Dummy_2003</td>
<td>-0.012</td>
<td>-0.024</td>
<td>-0.036</td>
<td>-0.025</td>
<td>-0.027</td>
<td>-0.013</td>
</tr>
<tr>
<td>Dummy_2004</td>
<td>-0.007</td>
<td>-0.022</td>
<td>-0.025</td>
<td>-0.020</td>
<td>-0.022</td>
<td>-0.009</td>
</tr>
<tr>
<td>Dummy_2005</td>
<td>-0.007</td>
<td>-0.025</td>
<td><strong>-0.004</strong></td>
<td>-0.029</td>
<td>-0.032</td>
<td>-0.013</td>
</tr>
<tr>
<td>Dummy_2006</td>
<td>-0.083</td>
<td>-0.059</td>
<td>0.054</td>
<td>0.049</td>
<td>-0.091</td>
<td>-0.120</td>
</tr>
<tr>
<td>Dummy_2007</td>
<td>-0.081</td>
<td>-0.058</td>
<td>0.091</td>
<td>0.062</td>
<td>-0.102</td>
<td>-0.128</td>
</tr>
<tr>
<td>Dummy_2008</td>
<td>-0.077</td>
<td>-0.053</td>
<td>0.119</td>
<td>0.103</td>
<td>-0.092</td>
<td>-0.134</td>
</tr>
<tr>
<td>Dummy_2009</td>
<td>-0.087</td>
<td>-0.062</td>
<td>0.129</td>
<td>0.107</td>
<td>-0.095</td>
<td>-0.135</td>
</tr>
<tr>
<td>Dummy_2010</td>
<td>-0.076</td>
<td>-0.056</td>
<td>0.140</td>
<td>0.115</td>
<td>-0.095</td>
<td>-0.138</td>
</tr>
<tr>
<td>Age</td>
<td>0.015</td>
<td>0.014</td>
<td>0.049</td>
<td>0.027</td>
<td>0.017</td>
<td>0.009</td>
</tr>
<tr>
<td>Age squared</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Children present</td>
<td><strong>0.001</strong></td>
<td><strong>-0.001</strong></td>
<td><strong>0.004</strong></td>
<td><strong>0.008</strong></td>
<td><strong>0.002</strong></td>
<td>0.010</td>
</tr>
<tr>
<td>Single</td>
<td>-0.011</td>
<td><strong>0.002</strong></td>
<td><strong>-0.002</strong></td>
<td><strong>-0.002</strong></td>
<td><strong>-0.012</strong></td>
<td>0.003</td>
</tr>
<tr>
<td>COB west</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.074</td>
<td>0.081</td>
<td>-</td>
</tr>
<tr>
<td>COB non-west</td>
<td><strong>0.002</strong></td>
<td>-</td>
<td>-0.045</td>
<td>-0.041</td>
<td>0.028</td>
<td>-0.112</td>
</tr>
<tr>
<td>COB Nordic</td>
<td>0.052</td>
<td>0.052</td>
<td>0.050</td>
<td>0.076</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COB w. Euro</td>
<td>0.029</td>
<td>0.057</td>
<td>0.057</td>
<td>0.058</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td><strong>0.004</strong></td>
<td><strong>0.004</strong></td>
<td><strong>0.004</strong></td>
<td><strong>0.004</strong></td>
<td><strong>0.004</strong></td>
<td><strong>0.004</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB other Euro</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COB Africa</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COB Lat. America</td>
<td><strong>0.034</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.069</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>-0.017</strong></td>
<td><strong>-0.027</strong></td>
</tr>
<tr>
<td>COB e. Asia n.dev.</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>0.019</strong></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB w. Africa</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.169</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YSA west</td>
<td><strong>0.000</strong></td>
<td><strong>0.000</strong></td>
<td>-0.002</td>
<td>-</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YSA non-west</td>
<td><strong>0.000</strong></td>
<td>-0.001</td>
<td>-0.002</td>
<td>-</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desc. non-west</td>
<td>-0.011</td>
<td>-0.011</td>
<td>-0.011</td>
<td>-0.011</td>
<td>-0.011</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desc. w. Euro</td>
<td>-</td>
<td>-</td>
<td><strong>-0.013</strong></td>
<td><strong>-0.029</strong></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desc. other Euro</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>-0.056</strong></td>
<td><strong>-0.080</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desc. Africa</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Att. educ. primary</td>
<td>-0.089</td>
<td>-0.040</td>
<td>-0.068</td>
<td>-0.094</td>
<td>-0.057</td>
<td>-0.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.048</td>
<td>-0.022</td>
</tr>
<tr>
<td>Att. ed. Secondary</td>
<td>-0.090</td>
<td>-0.085</td>
<td>-0.154</td>
<td>-0.143</td>
<td>-0.100</td>
<td>-0.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.048</td>
<td></td>
</tr>
<tr>
<td>Att. ed. Vocational</td>
<td>-0.130</td>
<td>-0.099</td>
<td>-0.131</td>
<td>-0.109</td>
<td>-0.154</td>
<td>-0.102</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.163</td>
<td>-0.225</td>
</tr>
<tr>
<td>Att. ed. tert. short</td>
<td><strong>-0.031</strong></td>
<td>-0.081</td>
<td>-0.034</td>
<td>-0.041</td>
<td>-0.100</td>
<td>-0.089</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.117</td>
<td>-0.108</td>
</tr>
<tr>
<td>Att. ed. tert. med.</td>
<td>-0.067</td>
<td>-0.059</td>
<td>-0.039</td>
<td>-0.041</td>
<td>-0.092</td>
<td>-0.090</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.118</td>
<td>-0.111</td>
</tr>
<tr>
<td>Att. ed. tert. long</td>
<td><strong>-0.055</strong></td>
<td><strong>-0.022</strong></td>
<td>-0.036</td>
<td>-0.041</td>
<td>-0.101</td>
<td>-0.097</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.105</td>
<td>-0.075</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. NA</td>
<td><strong>0.004</strong></td>
<td><strong>-0.002</strong></td>
<td>-0.006</td>
<td><strong>-0.006</strong></td>
<td>0.011</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>-0.010</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. 2</td>
<td><strong>-0.003</strong></td>
<td>0.024</td>
<td>0.024</td>
<td>0.029</td>
<td>0.011</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. 3</td>
<td>0.013</td>
<td><strong>0.002</strong></td>
<td>0.006</td>
<td>0.007</td>
<td>0.011</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. 4</td>
<td><strong>-0.001</strong></td>
<td><strong>-0.008</strong></td>
<td><strong>0.010</strong></td>
<td>0.011</td>
<td>0.011</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. 5</td>
<td>0.015</td>
<td><strong>-0.005</strong></td>
<td>0.017</td>
<td>0.013</td>
<td>0.011</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dad’s educ. 6</td>
<td>0.031</td>
<td><strong>0.008</strong></td>
<td>0.016</td>
<td>0.019</td>
<td>0.025</td>
<td><strong>0.002</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mum’s educ. NA</td>
<td><strong>0.004</strong></td>
<td><strong>0.002</strong></td>
<td>0.013</td>
<td><strong>-0.004</strong></td>
<td><strong>-0.024</strong></td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mum’s educ. 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><strong>-0.022</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mum’s educ. 2</td>
<td><strong>0.003</strong></td>
<td><strong>-0.002</strong></td>
<td>0.015</td>
<td>0.014</td>
<td>0.018</td>
<td>-0.026</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mum’s educ. 3</td>
<td>0.016</td>
<td>0.016</td>
<td>0.019</td>
<td>0.009</td>
<td>0.018</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mum’s educ. 4</td>
<td>0.035</td>
<td><strong>-0.014</strong></td>
<td>0.035</td>
<td>0.022</td>
<td>0.018</td>
<td>-0.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mum’s educ. 5</td>
<td><strong>0.000</strong></td>
<td>0.012</td>
<td>0.015</td>
<td><strong>0.003</strong></td>
<td>0.018</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The table contains statistical data with significance levels indicated by **0.001, 0.01, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.99.*
| Mum's educ. 6 | **-0.002** | **0.010** | **0.009** | 0.012 | **-0.006** | - | 0.032 | 0.053 | **0.000** | **0.001** | **0.007** | **0.016** |
| Experience | **0.001** | 0.009 | 0.012 | - | 0.007 | 0.006 | 0.005 | - | 0.010 | 0.006 | 0.003 | 0.002 |
| Exp. square root | 0.056 | - | 0.040 | 0.09 | 0.018 | 0.041 | 0.031 | 0.072 | 0.028 | 0.021 | 0.078 | 0.102 |
| Second. grade NA | - | - | 0.307 | 0.368 | 0.222 | 0.313 | 0.187 | 0.243 | 0.199 | 0.201 | 0.309 | 0.383 |
| Secondary grade | - | - | 0.045 | 0.053 | 0.046 | 0.053 | 0.031 | 0.037 | 0.032 | 0.032 | 0.037 | 0.045 |
| 1, UE benefits | -0.020 | -0.031 | -0.044 | -0.047 | -0.042 | -0.045 | -0.072 | -0.057 | -0.057 | -0.032 | -0.080 | -0.043 |
| 1, Soc. Sec. UE | -0.023 | -0.022 | -0.020 | -0.019 | -0.024 | -0.039 | -0.057 | -0.037 | -0.069 | -0.059 | -0.109 | **0.021** |
| 1, Soc. Sec. nUE | -0.028 | -0.027 | -0.020 | -0.017 | -0.049 | -0.015 | **-0.012** | -0.068 | -0.054 | -0.029 | -0.065 | **0.008** |
| 1, subs. empl. | -0.039 | -0.011 | -0.043 | -0.037 | -0.049 | -0.035 | - | -0.025 | -0.023 | -0.093 | -0.058 |
| 1, disab. benefits | **0.000** | - | -0.006 | **0.001** | -0.007 | 0.009 | -0.005 | 0.002 | 0.004 | 0.016 | **0.003** | 0.003 |
| 1, maternity leave | - | 0.006 | - | -0.034 | - | -0.029 | **-0.001** | -0.050 | -0.004 | -0.040 | **-0.002** | -0.048 |
| 1, other leave | **-0.011** | -0.024 | -0.047 | -0.086 | -0.036 | -0.033 | -0.054 | -0.065 | -0.081 | -0.078 | -0.087 | -0.107 |
| 1, not in lab. force | -0.009 | -0.013 | -0.022 | - | -0.036 | -0.037 | -0.015 | -0.037 | -0.034 | -0.055 | 0.018 | 0.016 |
| Wks, UE benefits | -0.134 | -0.116 | -0.195 | -0.130 | -0.193 | -0.114 | -0.179 | -0.146 | -0.184 | -0.091 | -0.216 | -0.153 |
| Wks, S. sec. UE | -0.050 | -0.079 | -0.165 | -0.062 | -0.137 | -0.089 | -0.233 | -0.155 | -0.101 | -0.071 | -0.291 | -0.149 |
| Wks, S. sec. nUE | -0.062 | -0.064 | -0.128 | -0.079 | -0.106 | -0.085 | -0.126 | -0.109 | -0.128 | **-0.042** | -0.209 | -0.210 |
| Wks, subs. empl. | -0.048 | -0.038 | 0.000 | - | -0.033 | -0.031 | -0.086 | -0.082 | -0.095 | - | -0.086 | -0.128 |
| Wks, disab. Ben. | -0.138 | -0.118 | 0.000 | -0.121 | -0.185 | -0.132 | -0.244 | -0.138 | -0.269 | -0.114 | -0.283 | -0.136 |
| Wks, mater. leave | - | - | - | - | - | -0.028 | **0.004** | - | **0.041** | -0.069 | **0.006** | -0.010 |
| Wks, other leave | -0.087 | **-0.006** | 0.000 | 0.021 | -0.108 | -0.045 | -0.067 | **0.000** | **0.009** | **0.004** | 0.059 | 0.036 |
| Wks, not in lab. f. | -0.123 | -0.066 | 0.000 | -0.099 | -0.160 | **-0.023** | -0.105 | -0.065 | -0.142 | **-0.022** | -0.164 | -0.055 |

Note: ** indicates that the estimate is not significant at a 5% level.
Appendix 2 Model estimation

The model is estimated by a Bayesian simulation method on the basis of iterative random draws from the parameters’ conditional distributions. In this appendix, the conditional posterior distributions for the model’s parameters are derived. The derivations of the conditional posterior distributions for the parameters \( \beta = (\beta_1, \beta_2), \gamma, \phi \) and \( \tau \) utilise the following differenced and \( \rho \)-conditional version of (1) and (2).

The model’s parameters and latent variables are grouped in nine blocks and a new set of values are drawn from the conditional posterior distributions for each iteration of the Gipps sampler:

(i) \( (\rho_{11}, \rho_{21}) \sim p((\rho_1, \rho_2) | \gamma_0, \beta_{10}, \beta_{20}, \phi_0, \tau_0, \sigma_{r_0}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(ii) \( \gamma_1 \sim p(\gamma | \rho_{11}, \rho_{21}, \beta_{10}, \beta_{20}, \phi_0, \tau_0, \sigma_{r_0}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(iii) \( (\beta_{11}, \beta_{21}) \sim p((\beta_1, \beta_2) | \rho_{11}, \rho_{21}, \gamma_1, \beta_{10}, \beta_{20}, \phi_0, \tau_0, \sigma_{r_0}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(iv) \( \phi_1 \sim p(\phi | \rho_{11}, \rho_{21}, \gamma_1, \beta_{11}, \beta_{21}, \phi_0, \tau_0, \sigma_{r_0}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(v) \( \tau_1 \sim p(\tau_1 | \rho_{11}, \rho_{21}, \gamma_1, \beta_{11}, \beta_{21}, \phi_0, \tau_0, \sigma_{r_0}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(vi) \( \sigma_{r_1}^2 \sim p(\sigma_{r_1}^2 | \rho_{11}, \rho_{21}, \gamma_1, \beta_{11}, \beta_{21}, \phi_0, \tau_1, \sigma_{r_0}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(vii) \( \sigma_{11}^2 \sim p(\sigma_{11}^2 | \rho_{11}, \rho_{21}, \gamma_1, \beta_{11}, \beta_{21}, \phi_0, \tau_1, \sigma_{r_1}^2, \sigma_{10}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(viii) \( \sigma_{21}^2 \sim p(\sigma_{21}^2 | \rho_{11}, \rho_{21}, \gamma_1, \beta_{11}, \beta_{21}, \phi_0, \tau_1, \sigma_{r_1}^2, \sigma_{11}^2, \sigma_{20}^2, Y_{20}, X, Y_1) \)

(ix) \( Y_{21} \sim p(Y_{21} | \rho_{11}, \rho_{21}, \gamma_1, \beta_{11}, \beta_{21}, \phi_0, \tau_1, \sigma_{r_1}^2, \sigma_{11}^2, \sigma_{21}^2, X, Y_1) \)

The model estimation and (2).

Prior distributions

The prior distributions are the following:

(i) \( \gamma \sim TN_{1-1,1}(\gamma, Y_2) \)
\( (\rho_1, \rho_2) \sim TN_{-1,1} \left( (\square, \square), \left[ \begin{array}{c} \nu \\ 0 \\ \nu \rho_2 \end{array} \right] \right) \)

\( (\beta_1, \beta_2) \sim N((\beta_1, \beta_2), V_{\beta}) \)

\( \phi \sim N(\phi, V_{\phi}) \)

\( f_i s_i^2 / \sigma^2 \sim \chi^2 (f_i) \) with \( f_i > 0 \) and \( s_i^2 > 0 \)

\( f_i s_i^2 / \sigma^2 \sim \chi^2 (f_i) \) with \( f_i > 0 \) and \( s_i^2 > 0 \)

\( f_i s_i^2 / \sigma^2 \sim \chi^2 (f_i) \) with \( f_i > 0 \) and \( s_i^2 > 0 \)

\( N(\mu, \Sigma) \) is a normal distribution with mean \( \mu \) and variance \( \Sigma \). \( TN_{a,b} (\mu, \Sigma) \) is a truncated normal distribution on the interval \( ]a,b[ \). \( \chi^2 (f) \) is a \( \chi^2 \)-distribution with \( f \) degrees of freedom. The parameters for the priors are chosen so that the posterior distributions are well-defined.

**Posterior distributions**

The presence of error term autocorrelation complicates the derivation of the posterior distributions. To overcome this, the first step derives the model in a differenced form, which conditional on \( \rho = (\rho_1, \rho_2) \) has no autocorrelation.

**The differenced model conditional on \( \rho = (\rho_1, \rho_2) \)**

Conditional on \( \rho_1 \) and \( \rho_2 \) the model is rewritten in its differenced form with unit variance error terms:

(A1) \( y_{it}^{\sigma_1} = x_{it}^{\sigma_1} \beta_1 + e_{i1} \sim N(0,1) \).

(A2) \( \tilde{y}_{i1}^{\sigma_2} = y_{i1}^{\sigma_2} + \tilde{x}_{i1}^{\sigma_2} \beta_2 - \rho_1 x_{i1}^{\sigma_2} \beta_1 + \tau_{i1}^{\sigma_2} + \phi e_{i1}^{\sigma_2} + e_{i2} = \sigma_2^{-1} \eta_{i2} \sim N(0,1) \).

(A3) \( \hat{y}_{i1}^{\sigma_2} = \tilde{y}_{i1}^{\sigma_2} + \tilde{x}_{i1}^{\sigma_2} \beta_2 + (1 - \rho_2) \tau_{i1}^{\sigma_2} + (1 - \rho_2) \phi e_{i1}^{\sigma_2} + e_{i2} \sim \text{nid}(0,1) \).

Where \( y_{i1}^{\sigma_1} = \sigma_1^{-1} y_{i1}, \tilde{y}_{i1}^{\sigma_2} = \sigma_2^{-1} (y_{i1} - \rho_1 y_{i1}) \) and \( \hat{y}_{i1}^{\sigma_2} = \sigma_2^{-1} (\tilde{y}_{i1} - \rho_2 \tilde{y}_{i1-1}) \).

The following result for the posterior distribution of a generalised linear regression model is utilised for most parameters.
Box 1 The posterior distribution of a generalized linear regression model

Conditional on the normal prior distribution $\beta \sim N(\beta, V_\beta)$, the posterior distribution of the generalised linear regression model $Y = X\beta + \varepsilon$ with heteroscedastic errors $\varepsilon \sim N(0, \Omega)$ and $\beta \sim N(\hat{\beta}, V_\beta)$ where $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$ and $V_\beta = (X'\Omega^{-1}X)^{-1}$ is given by $\beta \sim N(\bar{\beta}, V_\bar{\beta})$

where

$$V_\bar{\beta}^{-1} = V_\beta^{-1} + \bar{\varepsilon}^{-1}$$

and

$$\bar{\beta} = V_\beta^{-1} \bar{\beta} + V_\beta^{-1} \varepsilon$$

We are now ready to derive the posterior distributions for the model’s parameters.

**The posterior distribution for $\rho = (\rho_1, \rho_2)$**

The posterior distribution for $\rho = (\rho_1, \rho_2)$ is derived rearranging (3) and (4) in the following way:

**(3. $\rho$)**

$$\varepsilon_{i2} = \rho_1 \varepsilon_{i1} + \eta_{i2}$$

$$\varepsilon_{i2} = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \begin{pmatrix} \varepsilon_{i1} \\ 0 \end{pmatrix} + \eta_{i2}$$

**(4. $\rho$)**

$$\varepsilon_{it} = \rho_2 \varepsilon_{i,t-1} + \eta_{it}$$

$$\varepsilon_{it} = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \begin{pmatrix} \varepsilon_{i,t-1} \end{pmatrix} + \eta_{it}$$

By stacking these equations for every $i$

$$\begin{pmatrix} \varepsilon_{i2} \\ \varepsilon_{i3} \\ \vdots \\ \varepsilon_{it} \end{pmatrix} = \begin{pmatrix} \varepsilon_{i1} & 0 \\ 0 & \varepsilon_{i2} \\ \vdots & \vdots \\ 0 & \varepsilon_{i,t-1} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} + \begin{pmatrix} \eta_{i2} \\ \eta_{i3} \\ \vdots \\ \eta_{it} \end{pmatrix}$$

This is equation can be written as

$$\varepsilon_{i,2-T_i} = e_i^\rho \rho + \eta_{i,2-T_i}$$

Here $\eta_{i,2-T_i} \sim N(0, \Sigma_\rho)$ with $\Sigma_\rho = \sigma^2_{i,1} I_{T_i-1}$. By pre-multiplying by $\Sigma_\rho^{-1} = \sigma^{-1}_{i,1} I_{T_i-1}$ the above equation simplifies to

$$\Sigma_\rho^{-1} \varepsilon_{i,2-T_i} = \Sigma_\rho^{-1} e_i^\rho \rho + e_i^\rho$$

or

$$\Sigma_\rho^{-1} \varepsilon_{i,3-T_i} = \Sigma_\rho^{-1} e_i^\rho \rho + e_i^\rho$$

or

$$\Sigma_\rho^{-1} \varepsilon_{i,2-T_i} = \Sigma_\rho^{-1} e_i^\rho \rho + e_i^\rho$$

which can be stacked by $i$ to obtain:

$$Y^\rho = X^\rho + \varepsilon$$
This is another standard regression with white noise errors and
\[
\rho \sim N((X^\rho X^\rho)^{-1} X^\rho Y^\rho, (X^\rho X^\rho)^{-1})
\]
The posterior distribution for \( \rho = (\rho_1, \rho_2) \) is then derived by applying the result in Box 1.

**The posterior distribution for \( \beta = (\beta_1, \beta_2) \)**

The posterior distribution for \( \beta = (\beta_1, \beta_2) \) is derived by rearranging (A1)-(A3).

(A1. \( \beta \)) \[
y_{i1}^\beta = (x_{i1}^\beta 0) \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + e_{i1}
\]

(A2. \( \beta \)) \[
\dot{y}_{i1}^\beta - (\rho y_{i1}^\beta + \tau_i^\sigma + \phi \epsilon_{i1}^\sigma) = (- \rho x_{i1}^\sigma x_{i2}^\sigma x_{i2}^\beta \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}) + e_{i2}
\]

(A3. \( \beta \)) \[
\dot{y}_{i2}^\beta - (\rho y_{i2}^\beta + (1 - \rho_2) \tau_i^\sigma + (1 - \rho_2) \phi \epsilon_{i2}^\sigma) = \begin{pmatrix} x_{i1}^\sigma x_{i2}^\sigma \beta_i \\ \hat{x}_{i2}^\sigma \end{pmatrix} + e_{i2}
\]

These are stacked by \( t \) for every \( i : \)
\[
\begin{pmatrix}
y_{i1}^\beta \\
\dot{y}_{i1}^\beta - (\rho y_{i1}^\beta + \tau_i^\sigma + \phi \epsilon_{i1}^\sigma) \\
\dot{y}_{i2}^\beta - (\rho y_{i2}^\beta + (1 - \rho_2) \tau_i^\sigma + (1 - \rho_2) \phi \epsilon_{i2}^\sigma)
\end{pmatrix} = \begin{pmatrix} x_{i1}^\sigma x_{i2}^\sigma 0 \\ - \rho x_{i1}^\sigma x_{i2}^\sigma x_{i2}^\beta \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \\ \hat{x}_{i2}^\sigma \end{pmatrix} + e_{i2}
\]

The following compact form is then derived by further stacking the \( \beta \)-equation:
\[
Y^\beta = X^\beta + e
\]

This is a standard regression with white noise errors and hence
\[
\beta \sim N((X^\beta X^\beta)^{-1} X^\beta Y^\beta, (X^\beta X^\beta)^{-1})
\]
The posterior distribution for \( \beta \) is then derived by applying the result in Box 1.

**The posterior distribution for \( \gamma \)**

The posterior distribution for \( \gamma \) is derived in a similar manner by rearranging (A2) og (A3).

(A2. \( \gamma \)) \[
\dot{y}_{i1}^\gamma = \rho y_{i1}^\gamma + x_{i2}^\gamma \beta_2 - \rho x_{i1}^\gamma \beta_1 + \tau_i^\gamma + \phi \epsilon_{i1}^\gamma + e_{i2}
\]
\[
\dot{y}_{i2}^\gamma - (x_{i2}^\gamma \beta_2 - \rho x_{i1}^\gamma \beta_1 + \tau_i^\gamma + \phi \epsilon_{i2}^\gamma) = y_{i1}^\gamma \gamma + e_{i2}
\]
(A3. $\gamma$)
\[
\dot{y}_{i,t}^\gamma = \gamma_{i,t-1}^\gamma + \dot{x}_{i,t}^\gamma \beta_2 + (1-\rho_2)(\tau_{i,t}^\gamma + \phi e_{i,t}^\gamma) + e_{it}
\]
\[
\dot{y}_{i,t}^\gamma - (\dot{x}_{i,t}^\gamma \beta_2 + (1-\rho_2)(\tau_{i,t}^\gamma + \phi e_{i,t}^\gamma)) = \gamma_{i,t}^\gamma + e_{it}
\]

Stack again by $t$ for every $i$:
\[
\begin{bmatrix}
\dot{y}_{2,t}^\gamma - (x_{2,t}^\gamma \beta_2 - \rho_1 x_{1,t}^\gamma \beta_1 + \tau_{1,t}^\gamma + \phi e_{1,t}^\gamma) \\
\dot{y}_{2,t}^\gamma - (x_{2,t}^\gamma \beta_2 + (1-\rho_2)(\tau_{i,t}^\gamma + \phi e_{i,t}^\gamma)) \\
\vdots \\
\dot{y}_{2,t}^\gamma - (x_{2,t}^\gamma \beta_2 - \rho_1 x_{1,t}^\gamma \beta_1 + \tau_{1,t}^\gamma + \phi e_{1,t}^\gamma)
\end{bmatrix}
= \begin{bmatrix}
y_{1,t}^\gamma \\
y_{2,t}^\gamma \\
\vdots \\
y_{i,t}^\gamma \\
\dot{y}_{i-1,t}^\gamma
\end{bmatrix} + \begin{bmatrix}
y_{i,t}^\gamma \\
y_{i,t}^\gamma \\
\vdots \\
y_{i,t}^\gamma \\
\dot{y}_{i-1,t}^\gamma
\end{bmatrix}
\]

Then by further stacking by $i$:
\[
Y^\gamma = X^\gamma + e
\]

This is a standard regression with white noise errors and hence
\[
\gamma \sim N((X^\gamma, X^\gamma)^{-1} X^\gamma Y^\gamma, (X^\gamma, X^\gamma)^{-1})
\]

The posterior distribution for $\gamma$ is then expressed then derived by applying the result in Box 1.

The posterior distribution for $\phi = (\phi^1, \phi^2)$

The posterior distribution for $\phi$ is derived rearranging (A2) and (A3):

(A2. $\phi$)
\[
\dot{y}_{i,t}^\phi = \gamma_{i,t}^\phi + x_{i,t}^\gamma \beta_2 - \rho_1 x_{1,t}^\gamma \beta_1 + \tau_{i,t}^\gamma + \phi e_{i,t}^\gamma + e_{i,t}
\]
\[
\dot{y}_{i,t}^\phi - (x_{i,t}^\gamma \beta_2 - \rho_1 x_{1,t}^\gamma \beta_1 + \tau_{i,t}^\gamma + \phi e_{i,t}^\gamma) = \sigma_2^{-1} e_{i,t}
\]

(A3. $\phi$)
\[
\dot{y}_{i,t}^\phi = \gamma_{i,t-1}^\phi + \dot{x}_{i,t}^\gamma \beta_2 + (1-\rho_2)(\tau_{i,t}^\gamma + \phi e_{i,t}^\gamma) + e_{it}
\]
\[
\dot{y}_{i,t}^\phi - (\gamma_{i,t-1}^\phi + \dot{x}_{i,t}^\gamma \beta_2 + (1-\rho_2)(\tau_{i,t}^\gamma + \phi e_{i,t}^\gamma)) = \sigma_2^{-1} (1-\rho_2) e_{i,t}
\]

In stacked form, first for each $i$
\[
\begin{bmatrix}
\dot{y}_{2,t}^\gamma - (x_{2,t}^\gamma \beta_2 - \rho_1 x_{1,t}^\gamma \beta_1 + \tau_{1,t}^\gamma) \\
\dot{y}_{2,t}^\gamma - (x_{2,t}^\gamma \beta_2 + (1-\rho_2)(\tau_{i,t}^\gamma + \phi e_{i,t}^\gamma)) \\
\vdots \\
\dot{y}_{2,t}^\gamma - (x_{2,t}^\gamma \beta_2 - \rho_1 x_{1,t}^\gamma \beta_1 + \tau_{1,t}^\gamma)
\end{bmatrix}
= \begin{bmatrix}
\sigma_2^{-1} e_{i,t}
\sigma_2^{-1} (1-\rho_2) e_{i,t}
\vdots
\end{bmatrix}
\]

And then by $i$:
\[
Y^\phi = X^\phi \phi + e
\]
This is a standard regression with white noise errors and hence

\[ \phi \sim N\left( (X^\phi X^\phi)^{-1} X^\phi Y^\phi, (X^\phi X^\phi)^{-1} \right) \]

The posterior distribution for \( \phi \) is then derived by applying the result in Box 1.

**The posterior distribution for \( \sigma_2^2 \)**

The conditional posterior distribution for \( \sigma_2^2 \) is given by

\[ \left( \sum_{i=1}^{N} \tau_i^2 + \nu_\tau^2 \right) \sigma_2^2 \sim \chi^2(N + \nu_\tau) \]

\( \nu_\tau \) and \( \nu_\tau \) are the parameters (mean and degrees of freedom) for the prior distribution of \( \sigma_2^2 \).

**The posterior distribution for \( \tau_i \)**

The conditional posterior distribution for \( \tau_i \) is derived by first rearranging (A2) and (A3)

(A2. \( \tau_i \)) \[
\dot{y}_{i2}^\sigma = \gamma_{i1}^\sigma + x_{i2} \beta_2 - \rho_1 x_{i1}^\sigma \beta_1 + \tau_{i1}^\sigma + \phi e_{i1}^\sigma + e_{i2} \\
\dot{y}_{i2}^\sigma - (\gamma_{i1}^\sigma + x_{i2} \beta_2 - \rho_1 x_{i1}^\sigma \beta_1 + \phi e_{i1}^\sigma) = \sigma_2^{-1} \tau_i + e_{i2} 
\]

(A3. \( \tau_i \)) \[
\dot{y}_{i2}^\sigma = \gamma_{i2}^\sigma + x_{i2} \beta_2 + (1 - \rho_2) (\tau_{i1}^\sigma + \phi e_{i1}^\sigma) + e_{i2} \\
\dot{y}_{i2}^\sigma - (\gamma_{i2}^\sigma + x_{i2} \beta_2 + (1 - \rho_2) \phi e_{i1}^\sigma) = \sigma_2^{-1} (1 - \rho_2) \tau_i + e_{i2} 
\]

These equations are stacked for each \( i \):

\[
\begin{bmatrix}
\dot{y}_{i2}^\sigma - (\gamma_{i2}^\sigma + x_{i2} \beta_2 - \rho_1 x_{i1}^\sigma \beta_1 + \phi e_{i1}^\sigma)\\
\dot{y}_{i2}^\sigma - (\gamma_{i2}^\sigma + x_{i2} \beta_2 + (1 - \rho_2) \phi e_{i1}^\sigma)
\end{bmatrix} = \begin{bmatrix} \sigma_2^{-1} \\ \sigma_2^{-1} (1 - \rho_2) \end{bmatrix} \begin{bmatrix} \tau_i \\ e_{i2} \end{bmatrix} 
\]

An entry is added for the \( \sigma_2^2 \)-term in the likelihood function:

\[ 0 = \sigma_2^{-1} \tau_i + \sigma_2^{-1} \omega_i \]

Thereby arriving at the following regression for each \( i \):
\[
\begin{align*}
\begin{cases}
\hat{y}_{i2}^\tau - (\hat{\gamma}_{i2}^\tau + x_i^{\tau_2} \beta_2 - \rho_i x_i^{\tau_2} \beta_1 + \sigma_2^{-1} \phi_{i1}) \\
\hat{y}_{i1}^\tau - (\hat{\gamma}_{i1}^\tau + x_i^{\tau_2} \beta_2 + \sigma_2^{-1} (1 - \rho_2) \phi_{i1}) \\
\end{cases}
\begin{bmatrix}
\sigma_2^{-1} \\
\sigma_2^{-1} (1 - \rho_2) \\
\sigma^{-1} \\
\end{bmatrix}
\begin{bmatrix}
\tau_i \\
\epsilon_i \\
\end{bmatrix}
\end{align*}
\]

This can be written in the more compact form

\[ Y_i^\tau = X_i^\tau \tau_i + \epsilon_i \quad i = 1, \ldots, N \]

The conditional posterior distribution for \( \tau_i \) is given by

\[ \tau_i \sim N \left( (X_i^\tau, X_i^\tau)^{-1} X_i^\tau Y_i^\tau, (X_i^\tau, X_i^\tau)^{-1} \right) \quad i = 1, \ldots, N \]

At the end of each iteration, values for \( \sigma_1^2 \) and \( \sigma_2^2 \) are drawn from the following conditional posterior distributions

\[
\frac{1}{\sigma_1^2} = \frac{1}{\sigma_1^2} \sim \chi^2(N + \nu_1)
\]

\[
\frac{1}{\sigma_2^2} = \frac{1}{\sigma_2^2} \sim \chi^2(N + \nu_2)
\]

(\( \delta_1^2, \nu_1 \)) and (\( \delta_2^2, \nu_2 \)) are the prior distribution parameters (mean and degrees of freedom) for \( \sigma_1^2 \) and \( \sigma_2^2 \).

**Sampling the parameters in a mixed normal distribution**

Assume \{\( \tau_i \)\}_{i=1, \ldots, n} is an observed sample from a mixed normal distribution of \( m \) normal distributions \( n(\mu_j, \sigma_j^2) \) for \( j = 1, \ldots, m \), that is

\[ \tau_i = \sum_{j=1}^{m} e_{ij} (\alpha_j + \sigma_j \xi_i) \quad \xi_i \sim n(0,1) \]

The indicator matrix \( e = \{e_{ij}\}_{i=1, \ldots, n; j=1, \ldots, m} \) specifies the distributions for \{\( \tau_i \)\}_{i=1, \ldots, n}. \( e_{ij} = e_j \) for \( j = 1, \ldots, m \) are independent multinomial with \( P(e_{ij} = 1) = p_j \).

**Definitions:**
\[ \tau = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_n \end{bmatrix}, \quad e = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1m} \\ e_{21} & e_{22} & \cdots & e_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1} & e_{n2} & \cdots & e_{nm} \end{bmatrix} \]

\[ n_j = \sum_{i=1}^{n} e_{ij} \] specify the number of observations in each of the \( m \) groups.

The parameters for sampling are \( \alpha_j \), \( \sigma_j^2 \) and \( p_j \). The indicators \( e_{ij} \) are also sampled.

1. **\( \alpha_j \)** - the mean for the \( m \) distributions

Prior distribution: \( \alpha \sim n(\alpha, V_\alpha) \)

The posterior distribution (conditional on \( \sigma_j^2 \), \( p_j \) and \( e_{ij} \), cf. Box 1):

\[ \alpha \sim n(\bar{\alpha}, V_\alpha) \] where \( V_\alpha = (V_\alpha^{-1} + V_\alpha^{-1})^{-1} \) and \( \bar{\alpha} = V_\alpha(e^\tau \sigma + V_\alpha^{-1} \alpha) \)

Each element is shown in the following derivation.

The distribution for \( \{\tau_i\}_{i=1,\ldots,n} \) may be expressed as

\[ \tau = e\alpha + \varepsilon, \text{ where } \varepsilon \sim n(0, I \cdot V) \text{ and } V = \begin{bmatrix} \sigma_{g_1}^2 \\ \sigma_{g_2}^2 \\ \vdots \\ \sigma_{g_s}^2 \end{bmatrix}, \text{ where } g_i = \{ j : e_{ij} = 1 \} \]

This regression equation is normalized by element-multiplying (#) through by \( V^{-1/2} = (\sigma_{g_1}^{-1}, \sigma_{g_2}^{-1}, \ldots, \sigma_{g_s}^{-1})' \)

\[ \tau^\sigma = e^\sigma \alpha + \xi, \text{ where } \tau^\sigma = V^{-1/2} \# \tau, e^\sigma = V^{-1/2} \# e \text{ and } \xi \sim n(0, I) \]

This is a standard regression with parameters \( \alpha \) and independent standard normal error terms, which means that \( \alpha \sim n(\bar{\alpha}, V_\alpha) \) where \( \bar{\alpha} = (e^\tau \sigma)^{-1} e^\tau \sigma \) and \( V_\alpha = (e^\tau \sigma)^{-1} \). The above posterior follows immediately from the usual result for the posterior distribution of a generalized linear model.

2. **\( \sigma_j^2 \)** - the variance for the \( m \) distributions

Prior distribution: \( f_j S_j^2 / \sigma_j^2 \sim \chi^2(f_j) \) where \( S_j^2 > 0 \) and \( f_j^2 > 0 \)

The posterior distribution (conditional on \( \alpha_j \), \( p_j \) and \( e_{ij} \)):

\[ \frac{\left( \sum_{i=1}^{n} e_{ij}(\tau_i - \alpha_j)^2 + f_j S_j^2 \right) / \sigma_j^2}{\sigma_j^2} \sim \chi^2(n_j + f_j) \]
Computationally, $\sum_{i=1}^{n} e_{ij}(\tau_i - \alpha_j)^2$ can be efficiently expressed in matrix form as $e'(\tau - e\alpha)^{\#2}$. The operator $\#2$ means taking the square of each element in the vector $(\tau - e\alpha)$.

3. $\{p_j\}_{j=1,...,m}$ – the overall probabilities for each normal distribution.

Prior distribution: $p = (p_1,...,p_m) \sim Beta(r_1,...,r_m)$.

The posterior distribution (conditional on $\alpha_j$, $\sigma_j^2$ and $e_{ij}$):

$$ p \sim Beta(\{n_j + r_j\}_{j=1,...,m}) $$

4. The distribution of $e_{ij}$ conditional on $\alpha_j$, $\sigma_j^2$ and $p_j$

$e = \{e_{ij}\}_{i=1,...,n;j=1,...,m}$ are drawn from a multinomial distribution with parameters $p_{ij}$. These are proportional to

$$ P(e_{ij} = 1) \propto p_j \sigma_j^{-1} \exp \left[ -\frac{1}{2\sigma_j^2} (\tau_i - \alpha_j)^2 \right] $$

The $p_j$ do not add up to 1 (i.e. $\sum_{j=1}^{m} p_j \neq 1$) and hence require normalization prior to sampling of $\{e_{ij}\}_{i=1,...,n;j=1,...,m}$.

Computationally, $p_j \sigma_j^{-1} \exp \left[ -\frac{1}{2\sigma_j^2} (\tau_i - \alpha_j)^2 \right]$ can be efficiently expressed in matrix form as

$$ \exp \left[ -\frac{1}{2} (\tau \cdot 1_{nm} - 1_{nm} \alpha')^2 \sigma_d^{-2} \right] (\sigma_d^{-1} p_d) $$

Here the $\exp()$ and the power function $(\cdot)^2$ are element by element functions. $1_{nm}$ is an 1xm vector of 1 and $\sigma_d^{-2}$ (and $\sigma_d^{-1}$) are mxm diagonal matrices of $\sigma_j^{-2}$ (and $\sigma_j^{-1}$). Likewise $p_d$ is a diagonal matrix of the p-values.
Appendix 3 Simulating missing observations for the dependent variable

The missing observations for the dependent variable are treated as latent variables and simulated as an integrated step during the iterative estimation procedure. The observations are drawn from the conditional distribution at the end of each iteration, that is, conditional the model and its simulated parameters. The model is a dynamic panel model with complex intertemporal dependencies. Consequently, the missing observations are correlated with both past and future values of the dependent and independent variables. In addition, the situation is complicated by cases of consecutive missing observations that, because of the intertemporal dependencies, are correlated and must therefore be drawn from their joint distributions.

In the following, the joint conditional distributions for the missing observations are derived. The method is used for two purposes. Firstly, the joint conditional distributions are used for drawing the missing observations during the Bayes iterations. Secondly, conditional on the estimated model parameters values for the missing observations are simulated from the joint distributions conditional and thereby closing the ‘holes’ in the data.

The following result is used frequently $\sigma^2_e = \frac{\sigma^2_\epsilon}{1 - \rho^2}$ and $\text{Cov}(\epsilon_{it}, \epsilon_{i,t+1}) = \rho^2 \sigma^2_e = \frac{\rho^2 \sigma^2_\epsilon}{1 - \rho^2}$

The conditional joint distribution for consecutive missing observations is derived by apply extrapolation of the model’s main equation (2) in its forward and backward versions\(^9\)

(2f) \[ y_{it} = \gamma y_{i,t-1} + x_{it} \beta_2 + \tau_i^* + \epsilon_{it} \] with $\epsilon_{it} = \rho \epsilon_{i,t-1} + \epsilon_{it}^f$

(2b) \[ y_{it} = \gamma y_{i,t+1} + x_{it} \beta_2 + \tau_i^* + \epsilon_{it} \] with $\epsilon_{it} = \rho \epsilon_{i,t+1} + \epsilon_{it}^b$

With $h$ consecutive missing observations these extrapolations lead to $2h$ equations that, conditional on the model’s parameters, can be rearranged and stacked to form a simple GLS regression with known error term variance and with the missing observations as unknown parameters. The joint distribution of the missing observations is then multinomial normal with the usual mean and variance of the parameter estimates.

The details follow from the following compilations for the cases with up to four consecutive missing. Cases of five or more missing observations are relatively straightforward generalisations.

1. One missing observation: $y_{i,j}$ is missing, $y_{i,j-1}$ and $y_{i,j+1}$ are known

This is the simplest case. A straightforward rearrangement of (2f) and (2b)

\[
\begin{align*}
(y_{i,j} | t-1) & \\
\gamma y_{i,j-1} + x_{i,j} \beta_2 + \tau_i^* + \epsilon_{it} & = -\gamma y_{i,j-1} + x_{i,j} \beta_2 + \tau_i + \epsilon_{it}
\end{align*}
\]

\[ \tau_i^* = \tau_i + \phi \epsilon_{i,1} \]

\(^9\) For simplicity, the random effect and the period \(1 \ \tau_i = \tau_i + \phi \epsilon_{i,1} \)
\[-(\gamma_{i,t-1} + x_u \beta_2 + \tau_i^*) - \rho^* e_{it-1} = -y_{it} + e_{it}' \]

With \( \rho^* = \begin{cases} \rho_1 & hvis \ t = 1 \\ \rho_2 & hvis \ t \neq 1 \end{cases} \)

\( (y_{it}, \ t + 1) \)

\[ y_{it} = \gamma y_{i,t+1} + x_u \beta_2 + \tau_i^* + e_{it} \]

\[ -(\gamma y_{i,t+1} + (x_u \beta_2 + \tau_i^*)) = -y_{it} + e_{it} \]

\[ -(\gamma y_{i,t+1} + (x_u \beta_2 + \tau_i^*)) - \rho^* e_{it+1} = -y_{it} + e_{it}' \]

This is equivalent to the equation

\[
\begin{bmatrix}
 y_{i1}^1 \\
 y_{i2}^2
\end{bmatrix} =
\begin{bmatrix}
 -1 \\
 -1
\end{bmatrix} y_{it} +
\begin{bmatrix}
 e_{it}' \\
 e_{it}'
\end{bmatrix}
\]

or

\[ y_i = \Gamma_1 y_{it} + e_i \]

where

\[
\begin{bmatrix}
 y_{i1}^1 \\
 y_{i2}^2
\end{bmatrix} =
\begin{bmatrix}
 -(\gamma y_{i,t-1} + x_u \beta_2 + \tau_i^*) - \rho^* e_{it-1} \\
 -(\gamma y_{i,t+1} + (x_u \beta_2 + \tau_i^*)) - \rho^* e_{it+1}
\end{bmatrix}
\]

\[
-\begin{bmatrix}
 y_{i1}^1 \\
 y_{i2}^2
\end{bmatrix} -
\begin{bmatrix}
 x_u \beta_2 \\
 x_u \beta_2
\end{bmatrix} -\begin{bmatrix}
 1 \\
 1
\end{bmatrix} - \begin{bmatrix}
 \rho^* e_{it-1} \\
 \rho^* e_{it+1}
\end{bmatrix}
\]

and \( e_i \sim N(0, \Sigma) \), where \( \Sigma = \sigma_2^{-1} \begin{bmatrix}
 1 & 0 \\
 0 & 1
\end{bmatrix} \)

Multiplying by \( \Sigma_1^{-1/2} \) obtains:

\[ y_i^* = \Gamma_1^* y_{it} + e_i \]

This is a standard OLS with \( y_{it} \) as a parameter two observations and with a \( N(0,1) \) error term. The distribution for \( y_{it} \) is then

\[ y_{it} \sim N\left( (\Gamma_1^* \gamma_{i,t-1} \Gamma_1^*)^{-1} \Gamma_1^* y_{i,t-1}, (\Gamma_1^* \gamma_{i,t-1} \Gamma_1^*)^{-1} \right) \]

Notice that the variance matrix \( \Gamma_1^* \) does not depend on individual characteristics, and therefore only requires re-evaluation once per Gipps iteration.

2. Two missing observation:: \( y_{i,t} \) and \( y_{i,t+1} \) are missing, \( y_{i,t-1}, y_{i,t+2} \) are known

The case of two missing observations results in a regression with four observations, two from (2f) and two from (2b). First, by rearranging (2f) and to express \( y_{it} \) and \( y_{i,t+1} \) conditional on t-1:

\[ \sum_1^{-1/2} \] exists when \( \sum_1 \) positive definite, which is is normally the case. Nevertheless, standard GLS will lead to the same result.
\((y_{i,t} | t-1)\)
\[
y_{it} = \gamma y_{i,t-1} + x_{it} \beta_2 + \tau_i^* + \varepsilon_{it}
\]
\[-(\gamma y_{i,t-1} + x_{it} \beta_2 + \tau_i^*) = -y_{it} + \varepsilon_{it}
\]
\[-(\gamma y_{i,t-1} + x_{it} \beta_2 + \tau_i^*) - \rho^* \varepsilon_{it-1} = -y_{it} + \varepsilon_{it}^f
\]

\((y_{i,t+1} | t-1)\)
\[
y_{it+2} = \gamma(y y_{i,t} + x_{it+1} \beta_2 + \tau_{i}^* + \varepsilon_{it+1}) + x_{it+2} \beta_2 + \tau_i^* + \varepsilon_{it+2}
\]
\[-(\gamma^2 y_{i,t-1} + (\gamma x_{it} + x_{it+1} \beta_2 + (\gamma + 1) \tau_i^*) = -y_{it+1} + \gamma \varepsilon_{it} + \varepsilon_{it+1}
\]
\[-(\gamma^2 y_{i,t-1} + (\gamma x_{it} + x_{it+1} \beta_2 + (\gamma + 1) \tau_i^*) - \rho^* (\gamma + \rho_2) \varepsilon_{it-1} =
\]
\[-y_{it+1} + (\gamma + \rho_2) \varepsilon_{it-1} + \varepsilon_{it+1}^f
\]

Second, by rearranging (2b) and to express \(y_{i,t}\) and \(y_{i,t+1}\) conditional on \(t+2\):
\[(y_{i,t+1} | t+2)\]
\[
y_{it+1} = \gamma y_{i,t+2} + x_{it+1} \beta_2 + \tau_i^* + \varepsilon_{it+1}
\]
\[-(\gamma y_{i,t+2} + x_{it+1} \beta_2 + \tau_{i}^*) - \rho_2 \varepsilon_{it+2} = -y_{it+1} + \varepsilon_{it+1}^b
\]
\[(y_{i,t} | t+2)\]
\[
y_{it} = \gamma y_{i,t+2} + x_{it} \beta_2 + \tau_i^* + \varepsilon_{it}
\]
\[
y_{it} = \gamma^2 y_{i,t+2} + (\gamma x_{it+1} + x_{it} \beta_2 + (\gamma + 1) \tau_i^*) + \gamma \varepsilon_{it+1} + \varepsilon_{it}
\]
\[-(\gamma^2 y_{i,t+2} + (\gamma x_{it+1} + x_{it} \beta_2 + (\gamma + 1) \tau_i^*) - (\gamma \rho_2 + \rho_2^2) \varepsilon_{it+2} =
\]
\[-y_{it} + (\gamma + \rho_2) \varepsilon_{it+1} + \varepsilon_{it}^b
\]

This is equivalent to the following regression with \(y_{i,t}\) and \(y_{i,t+1}\) as parameters:

\[
\begin{bmatrix}
y_{1t} \\
y_{2t} \\
y_{3t} \\
y_{4t}
\end{bmatrix}
= \begin{bmatrix}
-1 & 0 \\
0 & -1 \\
0 & -1 \\
-1 & 0
\end{bmatrix}
\begin{bmatrix}
y_{it} \\
y_{it+1}
\end{bmatrix}
+ \begin{bmatrix}
e_{it}^f \\
(\gamma + \rho_2) e_{it+1}^b + \varepsilon_{it+1}^b
\end{bmatrix}
\]

or
\[
y_i = \Gamma_2 \begin{bmatrix}
y_{it} \\
y_{it+1}
\end{bmatrix}
+ e_i
\]

where
\[
y_i = \begin{bmatrix}
y_{1t} \\
y_{2t} \\
y_{3t} \\
y_{4t}
\end{bmatrix}
= \begin{bmatrix}
-(\gamma y_{i,t+1} + x_{it} \beta_2 + \tau_i^*) - \rho^* \varepsilon_{it+1} \\
-(\gamma^2 y_{i,t-1} + (\gamma x_{it} + x_{it+1} \beta_2 + (\gamma + 1) \tau_i^*) - (\gamma \rho_2 + \rho_2^2) \varepsilon_{it+2}
\end{bmatrix}
\]
\[
\begin{bmatrix}
\gamma_1 y_{i,t-1} \\
\gamma_2 y_{i,t-1} \\
\gamma_1 y_{i,t+2} \\
\gamma_2 y_{i,t+2}
\end{bmatrix} - 
\begin{bmatrix}
x_t \\
x_{t+1} \\
x_t \\
x_{t+1}
\end{bmatrix} 
= 
\begin{bmatrix}
0 \\
0 \\
1 \\
1
\end{bmatrix} 
+ 
\begin{bmatrix}
-\rho^* e_{it-1} \\
-\rho^* (\gamma + \rho_2^*) e_{it-1} \\
-\rho_2^* e_{it+1} \\
-\rho_2^* (\gamma + \rho_2^*) e_{it+1}
\end{bmatrix}
\]

and \( e_i \sim N(0, \Sigma_2) \), where

\[
\Sigma_2 = \sigma_2^2 \begin{bmatrix}
1 & (\gamma + \rho_2^*) & 0 & 0 \\
(\gamma + \rho_2^*) & (\gamma + \rho_2^*)^2 + 1 & 0 & 0 \\
0 & 0 & 1 & (\gamma + \rho_2^*) \\
0 & 0 & (\gamma + \rho_2^*) & (\gamma + \rho_2^*)^2 + 1
\end{bmatrix}
\]

Multiplying by \( \Sigma_2^{-1/2} \) to get

\[
y_i^* = \Gamma_2^{1/2} y_{it+1} + e_i
\]

This is a standard OLS with four observations and \( N(0,1) \) error terms. The distribution for the parameters \( y_{it} \) and \( y_{i,t+1} \) is

\[(y_{it}, y_{it+1}) \sim N((\Gamma_2^{1/2})^{-1} \Gamma_2^{1/2} y_i^*, (\Gamma_2^{1/2})^{-1})\]

3. Three missing observation: \( y_{i,t-1}, y_{i,t+1} \) and \( y_{i,t+2} \) are missing, \( y_{i,t-1} \) log, \( y_{i,t+3} \) are known

The case of three missing observations results in a regression with six observations. First, by rearranging (2f) and to express \( y_{i,t-1}, y_{i,t+1} \) and \( y_{i,t+2} \) conditional on \( t-1 \):

\[(y_{i,t-1} | t-1) \]

\[
y_{it} = \gamma y_{i,t-1} + x_t \beta_2 + \tau_i^* + e_{it}
\]

\[-(\gamma y_{i,t-1} + x_t \beta_2 + \tau_i^*) - \rho^* e_{it-1} = -y_{it} + e_i^{f}
\]

\[(y_{i,t+1} | t-1) \]

\[
y_{it+1} = \gamma y_{i,t-1} + x_{it+1} \beta_2 + \tau_i^* + e_{it+1}
\]

\[-(\gamma^2 y_{i,t-1} + (x_t + x_{it+1}) \beta_2 + (\gamma + 1) \tau_i^*) - \rho^* (\gamma + \rho_2^*) e_{it-1} = -y_{it+1} + (\gamma + \rho_2^*) e_i^{f} + e_i^{f1}
\]

\[(y_{i,t+2} | t-1) \]

\[
y_{it+2} = \gamma (y_{i,t-1} + x_{it+1} \beta_2 + \tau_i^* + e_{it+1}) + x_{it+2} \beta_2 + \tau_i^* + e_{it+2}
\]

\[-(\gamma^3 y_{i,t-1} + (\gamma^2 x_t + x_{it+1} + x_{it+2}) \beta_2 + (\gamma^2 + \gamma + 1) \tau_i^*) - \rho^* (\gamma^2 + \gamma \rho_2 + \rho_2^2) e_{it-1} = -y_{it+2} + (\gamma^2 + \gamma \rho_2 + \rho_2^2) e_i^{f} + (\gamma + \rho_2^*) e_i^{f1} + e_i^{f2}
\]

And then, by rearranging (2b) and to express the missing observations conditional on \( t+2 \):
\[(y_{i,t+2} | t+3) \quad y_{it+2} = \gamma y_{i,t+3} + x_{it+2} \beta_2 + \tau_i^* + \epsilon_{it+2} - (\gamma y_{i,t+3} + x_{it+2} \beta_2 + \tau_i^*) - \rho \epsilon_{it+3} = -y_{it+2} + \epsilon_{it+2}^b\]

\[(y_{i,t+1} | t+3) \quad y_{it+1} = \gamma y_{i,t+2} + x_{it+1} \beta_2 + \tau_i^* + \epsilon_{it+1} - (\gamma^2 y_{i,t+2} + x_{it+1} \beta_2 + (\gamma + 1) \tau_i^*) - (\gamma \rho_2 + \rho_2^2) \epsilon_{it+3} = -y_{it+1} + (\gamma + \rho_2) \epsilon_{it+2}^b + \epsilon_{it+1}^b\]

\[(y_{i,t} | t+3) \quad y_{it} = \gamma y_{i,t+1} + x_{it} \beta_2 + \tau_i^* + \epsilon_{it} - (\gamma^3 y_{i,t+2} + x_{it+1} \beta_2 + (\gamma^2 + \gamma + 1) \tau_i^*) - (\gamma^2 \rho_2 + \gamma \rho_2^2 + \rho_2^3) \epsilon_{it+3} = -y_{it} + (\gamma^2 + \gamma \rho_2 + \rho_2^2) \epsilon_{it+2}^b + (\gamma + \rho_2) \epsilon_{it+1}^b + \epsilon_{it}^b\]

This is equivalent to the following regression with \(y_{i,t}, y_{i,t+1}\) and \(y_{i,t+2}\) as parameters:

\[
\begin{pmatrix}
    y_i^1 \\
    y_i^2 \\
    y_i^3 \\
    y_i^4 \\
    y_i^5 \\
    y_i^6
\end{pmatrix}
= \begin{pmatrix}
    -1 & 0 & 0 \\
    0 & -1 & 0 \\
    0 & 0 & -1 \\
    0 & 0 & -1 \\
    0 & -1 & 0 \\
    -1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
    y_{it} \\
    y_{it+1} \\
    y_{it+2}
\end{pmatrix}
+ \begin{pmatrix}
    e_{it}^f \\
    (\gamma + \rho_2) e_{it+1}^f + \epsilon_{it+1}^b \\
    (\gamma^2 + \gamma \rho_2 + \rho_2^2) e_{it+2}^f + (\gamma + \rho_2) e_{it+1}^b + \epsilon_{it+1}^b \\
    (\gamma^2 + \gamma \rho_2 + \rho_2^2) e_{it+2}^b + (\gamma + \rho_2) e_{it+1}^b + \epsilon_{it+1}^b
\end{pmatrix}
\]

or

\[
y_i = \Gamma_3 \begin{pmatrix}
y_{it} \\
y_{it+1} \\
y_{it+2}
\end{pmatrix} + \epsilon_i
\]

where

\[
y_i = \begin{pmatrix}
    - (\gamma y_{i,t-1} + x_{it-1} \beta_2 + \tau_i^*) - \rho^* \epsilon_{i,t-1} \\
    - (\gamma^2 y_{i,t-1} + (x_{it-1} + x_{it+1}) \beta_2 + (\gamma + 1) \tau_i^*) - \rho^* (\gamma + \rho_2) \epsilon_{i,t-1} \\
    - (\gamma^3 y_{i,t-1} + (x_{it+1} x_{it+2} + x_{it+2}) \beta_2 + (\gamma^2 + \gamma + 1) \tau_i^*) - \rho^* (\gamma^2 + \gamma \rho_2 + \rho_2^3) \epsilon_{i,t-1} \\
    - (\gamma y_{i,t+1} + x_{it+2} \beta_2 + \tau_i^*) - \rho \epsilon_{i,t+3} \\
    - (\gamma^2 y_{i,t+1} + (x_{it+2} + x_{it+1}) \beta_2 + (\gamma + 1) \tau_i^*) - (\gamma \rho_2 + \rho_2^2) \epsilon_{i,t+3} \\
    - (\gamma^3 y_{i,t+1} + (x_{it+2} + x_{it+1} + x_{it+2}) \beta_2 + (\gamma^2 + \gamma + 1) \tau_i^*) - (\gamma^2 \rho_2 + \gamma \rho_2^2 + \rho_2^3) \epsilon_{i,t+3}
\end{pmatrix}
\]
\[ \begin{align*}
\mathbf{Y}_{i,t+1} & = \begin{bmatrix} \gamma^2 \mathbf{Y}_{i,t-1} \\
\gamma^3 \mathbf{Y}_{i,t-3} \\
\gamma^3 \mathbf{Y}_{i,t+3} \\
\end{bmatrix} - \gamma \begin{bmatrix} x_{it} \\
x_{it+2} \\
x_{it+4} \\
\end{bmatrix} \begin{bmatrix} 0 \\
0 \\
0 \\
\end{bmatrix} \begin{bmatrix} 0 \\
0 \\
0 \\
\end{bmatrix} \begin{bmatrix} \beta_2 - \tau_i \gamma^2 + \gamma + 1 \\
\gamma^2 + \gamma + 1 \\
\gamma^2 + \gamma + 1 \\
\end{bmatrix} \begin{bmatrix} \rho^* e_{u-1} \\
\rho^* (\gamma + \rho_2) e_{u-1} \\
\rho^* (\gamma^2 + \gamma \rho_2 + \rho_2^2) e_{u-1} \\
\rho e_{u+3} \\
(\gamma \rho_2 + \rho_2^2) e_{u+3} \\
(\gamma^2 \rho_2 + \gamma \rho_2^2 + \rho_2^3) e_{u+3} \\
\end{bmatrix}
\end{align*} \]

and \( e_i \sim N(0, \Sigma_3) \), with \( \Sigma_3 = \sigma_3^2 \begin{bmatrix} \Sigma_3^{11} & 0 \\
0 & \Sigma_3^{13} \\
\end{bmatrix} \), where

\[ \Sigma_3^{11} = \begin{bmatrix}
1 & (\gamma + \rho_2) \\
(\gamma + \rho_2) & (\gamma + \rho_2)^2 + 1 \\
(\gamma^2 + \gamma \rho_2 + \rho_2^2) & (\gamma + \rho_2)(\gamma^2 + \gamma \rho_2 + \rho_2^2) + (\gamma + \rho_2) \\
(\gamma^2 + \gamma \rho_2 + \rho_2^2) & (\gamma + \rho_2)(\gamma^2 + \gamma \rho_2 + \rho_2^2) + (\gamma + \rho_2) \\
\end{bmatrix} \]

For computational simplicity, this matrix is decomposed as

\[ \Sigma_3 = \Sigma_3^{13} \Sigma_3^{13T} \]

where

\[ \Sigma_3^{13} = \begin{bmatrix}
1 & 0 \\
(\gamma + \rho_2) & 1 \\
(\gamma^2 + \gamma \rho_2 + \rho_2^2) & (\gamma + \rho_2) \\
\end{bmatrix} \]

As usual pre-multiply by \( \Sigma_3^{-1/2} \) to obtain

\[ \mathbf{y}_i = \Gamma_3 \begin{bmatrix} y_{i,t} \\
y_{i,t+1} \\
y_{i,t+2} \\
\end{bmatrix} + e_i \]

And again, this is a standard OLS with six observations and \( N(0,1) \) error terms. The distribution for the parameters \( y_{i,t}, y_{i,t+1} \) and \( y_{i,t+2} \) is

\[ (y_{i,t}, y_{i,t+1}, y_{i,t+2}) \sim N \left( \left( \Gamma_3^{-1} \right)^T \Gamma_3 \mathbf{y}_i, (\Gamma_3^{-1} \Gamma_3^{-1})^{-1} \right) \]

4. Four missing observations: \( y_{i,t}, x_{i,t+1}, y_{i,t+2} \) and \( y_{i,t+3} \) are missing, \( y_{i,t-1} \) and \( y_{i,t+3} \) are known

The case of three missing observations results in a regression with eight observations. First, by rearranging (2f) and to express \( y_{i,t}, y_{i,t+1}, y_{i,t+2} \) conditional on \( t-1 \):

\[ (y_{i,t} \mid t-1) \]

\[ y_{it} = \gamma y_{i,t-1} + x_{it} \beta_2 + \tau_i + e_{it} \]

\[ y_{it} = \gamma y_{i,t-1} + x_{it} \beta_2 + \tau_i + e_{it} \]

\[ (y_{i,t} \mid t-1) \]

\[ y_{i,t+1} = \gamma y_{i,t} + x_{i,t+1} \beta_2 + \tau_i + e_{i,t+1} \]

\[ y_{i,t+1} = \gamma (y_{i,t} + x_{i,t+1} \beta_2 + \tau_i + e_{i,t}) + x_{i,t+1} \beta_2 + \tau_i + e_{i,t+1} \]
\[-(y^2y_{i,t-1} + (\gamma x_t + x_{i,t-1})\beta_2 + (\gamma + 1)t_i^*) - \rho^*(\gamma + \rho_2)\epsilon_{it-1} = \]
\[-y_{it+1} + (\gamma + \rho_2)(e_{it} + e_{it}^f)\]

\[(y_{i,t+2} | t-1) y_{it+2} = \gamma(y_{i,t} + x_{it+1}\beta_2 + t_i^* + \epsilon_{it+1}) + x_{it+2}\beta_2 + \tau_i^* + \epsilon_{it+2}\]
\[-(y^3y_{i,t-1} + (y^2x_t + \gamma x_{it+1} + x_{i,t+2})\beta_2 + (y^2 + \gamma + 1)t_i^*)) - \rho^*(y^2 + \gamma \rho_2 + \rho_2^2)\epsilon_{it-1} = \]
\[-y_{it+2} + (y^2 + \gamma \rho_2 + \rho_2^2)e_{it}^f + (y + \rho_2)e_{it+1}^f + e_{it+2}^f\]

\[(y_{i,t+3} | t-1) y_{it+3} = \gamma(y_{i,t+1} + x_{it+2}\beta_2 + t_i^* + \epsilon_{it+2}) + x_{it+3}\beta_2 + \tau_i^* + \epsilon_{it+3}\]
\[-(y^4y_{i,t-1} + (y^3x_t + y^2x_{it+1} + \gamma x_{it+2} + x_{it+3})\beta_2 + (y^3 + y^2 + \gamma + 1)t_i^*)) - \rho^*(y^3 + \gamma \rho_2^0 + \gamma \rho_2^2 + \rho_2^3)\epsilon_{it-1} = \]
\[-y_{it+3} + (y^3 + \gamma \rho_2 + \gamma \rho_2^2 + \rho_2^3)e_{it}^f + (y^2 + \gamma \rho_2 + \rho_2^2)e_{it+1}^f + (y + \rho_2)e_{it+2}^f + e_{it+3}^f\]

Rearranging (2b) to express the missing observations conditional on \(t+2\):

\[(y_{i,t+3} | t+4) y_{it+3} = \gamma y_{i,t+4} + x_{it+3}\beta_2 + \tau_i^* + \epsilon_{it+3}\]
\[-(\gamma y_{i,t+4} + x_{it+3}\beta_2 + \tau_i^*) - \rho e_{it+4} = -y_{it+3} + e_{it+3}^b\]

\[(y_{i,t+2} | t+4) y_{it+2} = \gamma y_{i,t+3} + x_{it+2}\beta_2 + \tau_i^* + \epsilon_{it+2}\]
\[-(y^2y_{i,t+4} + (\gamma x_{it+3} + x_{it+2})\beta_2 + (\gamma + 1)t_i^*) - (\gamma \rho_2 + \rho_2^2)e_{it+4} = \]
\[-y_{it+2} + (\gamma + \rho_2)e_{it+3}^b + e_{it+2}^b\]

\[(y_{i,t+1} | t+4) y_{it+1} = \gamma y_{i,t+2} + x_{it+1}\beta_2 + \tau_i^* + \epsilon_{it+1}\]
\[-(\gamma^3y_{i,t+4} + (y^2x_{it+3} + \gamma x_{it+2} + x_{it+1})\beta_2 + (y^2 + \gamma + 1)t_i^*) - (y^2\rho_2 + y\rho_2^2 + \rho_2^3)e_{it+4} = \]
\[-y_{it+1} + (y^2 + \gamma \rho_2 + \rho_2^2)e_{it+3}^b + (y + \rho_2)e_{it+2}^b + e_{it+1}^b\]

\[(y_{i,t} | t+4) y_{it} = \gamma y_{i,t+1} + x_t\beta_2 + \tau_i^* + \epsilon_{it}\]
\[-(\gamma^4y_{i,t+4} + (y^3x_{it+3} + y^2x_{it+2} + \gamma x_{it+1} + x_t)\beta_2 + (y^3 + y^2 + \gamma + 1)t_i^*) - (y^3\rho_2 + y^2\rho_2^2 + y\rho_2^3 + \rho_2^4)e_{it+4} = \]
\[-y_{it} + (y^3 + \gamma \rho_2 + \gamma \rho_2^2 + \rho_2^3)e_{it+3}^b + (y^2 + \gamma \rho_2 + \rho_2^2)e_{it+2}^b + (y + \rho_2)e_{it+1}^b + e_{it}^b\]

This is equivalent to the following regression with \(y_{i,t}, y_{i,t+1}, y_{i,t+2}\) and \(y_{i,t+3}\) as parameters:
\[
\begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
y_6 \\
y_7 \\
y_8 
\end{bmatrix} =
\begin{bmatrix}
-1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1 \\
0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 
\end{bmatrix}
+ \begin{bmatrix}
y_{it} \\
y_{it+1} \\
y_{it+2} \\
y_{it+3} 
\end{bmatrix}
\]

or
\[
y_j = \Gamma_4 \begin{bmatrix}
y_{it} \\
y_{it+1} \\
y_{it+2} \\
y_{it+3} 
\end{bmatrix} + e_i
\]

where
\[
\begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
y_6 \\
y_7 \\
y_8 
\end{bmatrix} =
\begin{bmatrix}
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} \\
-(y_{i,t-1} + x_{it} + \beta_2 + \tau_i^*) - \rho_2 e_{it-1} 
\end{bmatrix}
\]

\[
\begin{bmatrix}
e_{it} \\
e_{it+1} \\
e_{it+2} \\
e_{it+3} 
\end{bmatrix} =
\begin{bmatrix}
(y + \rho_2) e_{it} + e_{it+1} \\
(y^2 + \rho_2 + \rho_2^2) e_{it} + (y + \rho_2) e_{it+1} + e_{it+2} \\
(y^3 + \rho_2 + \rho_2^2 + \rho_2^3) e_{it} + (y^2 + \rho_2 + \rho_2^2) e_{it+1} + (y + \rho_2) e_{it+2} + e_{it+3} \\
(y + \rho_2) e_{it+3} + e_{it+2} \\
(y + \rho_2) e_{it+3} + e_{it+2} \\
(y^2 + \rho_2 + \rho_2^2) e_{it+3} + (y + \rho_2) e_{it+2} + e_{it+1} \\
(y^3 + \rho_2 + \rho_2^2 + \rho_2^3) e_{it+3} + (y^2 + \rho_2 + \rho_2^2) e_{it+2} + e_{it+1} \\
(y + \rho_2) e_{it+3} + e_{it+2} 
\end{bmatrix}
\]
\[
\begin{align*}
&\begin{pmatrix}
\gamma y_{i,t-1} \\
\gamma^2 y_{i,t-1} \\
\gamma^3 y_{i,t-1} \\
\gamma^4 y_{i,t-1} \\
\gamma^2 y_{i,t+4} \\
\gamma^3 y_{i,t+4} \\
\gamma^4 y_{i,t+4}
\end{pmatrix} =
\begin{pmatrix}
x_{it} \\
x_{it+1} \\
x_{it+2} \\
x_{it+3} \\
x_{it} \\
x_{it+1} \\
x_{it+2}
\end{pmatrix} - \begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix} - \begin{pmatrix}
\gamma^2 \\
\gamma^3 \\
\gamma^4
\end{pmatrix} \begin{pmatrix}
x_{it} \\
x_{it+1} \\
x_{it+2} \\
x_{it+3}
\end{pmatrix} + \begin{pmatrix}
\beta_2
\end{pmatrix}
\end{align*}
\]

and \( \varepsilon_i \sim N(0, \Sigma_4) \) with \( \Sigma_4 = \sigma_2^2 \begin{pmatrix} \Sigma_1^n & 0 \\ 0 & \Sigma_1^n \end{pmatrix} \) where

\[
\Sigma_1^n = \begin{pmatrix}
1 & \gamma + \rho_2 \\
\gamma + \rho_2 & (\gamma + \rho_2)^2 + 1 \\
\gamma^2 + \gamma \rho_2 + \rho_2^2 & (\gamma + \rho_2)(\gamma^2 + \gamma \rho_2 + \rho_2^2) + (\gamma + \rho_2) \\
\gamma^3 + \gamma^2 \rho_2 + \gamma \rho_2^2 + \rho_2^3 & (\gamma + \rho_2)(\gamma^3 + \gamma^2 \rho_2 + \gamma \rho_2^2 + \rho_2^3) + (\gamma^2 + \gamma \rho_2 + \rho_2^2)
\end{pmatrix}
\]

For computational simplicity, this matrix is decomposed as

\[
\Sigma_1^n = \Sigma_1^{11} \Sigma_4^{11}
\]

where
\[
\Sigma_4^{-1/2} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
\gamma + \rho_2 & 1 & 0 & 0 \\
\gamma^2 + \gamma \rho_2 + \rho_2^2 & \gamma + \rho_2 & 1 & 0 \\
\gamma^3 + \gamma^2 \rho_2 + \gamma \rho_2^2 + \rho_2^3 & \gamma^2 + \gamma \rho_2 + \rho_2^2 & \gamma + \rho_2 & 1
\end{pmatrix}
\]

Pre-multiply by \(\Sigma_4^{-1/2}\) to obtain

\[
y_i^r = \Gamma_4^{\Sigma} \begin{pmatrix}
y_{it} \\
y_{it+1} \\
y_{it+2} \\
y_{it+3}
\end{pmatrix} + e_i
\]

This is a standard OLS with six observations and N(0,1) error terms where the joint distribution for the parameters \(y_{it}, y_{it+1}, y_{it+2}\) and \(y_{it+3}\) is given by

\[
(y_{it}, y_{it+1}, y_{it+2}) \sim N\left(\Gamma_4^{\Sigma} y_i^r, (\Gamma_4^{\Sigma} \Gamma_4^{\Sigma})^{-1}\right)
\]