Macroeconometrics and time series analysis

Edited by

Steven N. Durlauf and Lawrence E. Blume

THE NEW PALGRAVE

Macroeconometrics and Time Series Analysis

Edited by
Steven N. Durlauf
University of Wisconsin-Madison, USA

Lawrence E. Blume
Cornell University, USA

palgrave
macmillan

All articles first published in *The New Palgrave Dictionary of Economics*, 2nd Edition
Edited by Steven N. Durlauf and Lawrence E. Blume
in eight volumes, 2008

First published 2010 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

ISBN 978-0-230-23884-8 hardback
ISBN 978-0-230-23885-5 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

Printed and bound in Great Britain by
CPI Antony Rowe, Chippenham and Eastbourne

List of Contributors

LAURENT E. CALVET
HEC School of Management, France

TIMOTHY COGLEY
New York University, USA

FRANCIS X. DIEBOLD
University of Pennsylvania, USA

JESÚS FERNÁNDEZ-VILLAVERDE
University of Pennsylvania, USA

A. RONALD GALLANT
Duke University, USA

CLIVE W. J. GRANGER
University of California San Diego, USA

JAMES D. HAMILTON
University of California San Diego, USA

LARS PETER HANSEN
University of Chicago, USA

ANDREW HARVEY
University of Cambridge, UK

DAVID F. HENDRY
University of Oxford, UK

YONGMIAO HONG
Cornell University, USA

SVEND HYLLEBERG
Aarhus University, Denmark

LUTZ KILIAN
University of Michigan, USA

ANDROS KOURTELLOS
University of Cyprus, Cyprus

G. M. KUERSTEINER
University of California Davis, USA

KURT F. LEWIS
University of Iowa, USA

OLIVER B. LINTON
London School of Economics, UK

HELMUT LÜTKEPOHL
European University Institute, Italy

BRUCE MIZRACH
Rutgers University, USA

CHARLES R. NELSON
University of Washington, USA

MARC NERLOVE
University of Maryland, USA

PIERRE PERRON
Boston University, USA

PETER C. B. PHILLIPS
Yale University, USA

WERNER PLOBERGER
Washington University in St Louis, USA

DALE J. POIRIER
University of California Irvine, USA

JAMES B. RAMSEY
New York University, USA

P. M. ROBINSON
London School of Economics, UK

JUAN F. RUBIO-RAMÍREZ
Duke University, USA

THOMAS J. SARGENT
New York University, USA

FRANK SCHORFHEIDE
University of Pennsylvania, USA

NEIL SHEPHARD
University of Oxford, UK

CHRISTOPHER A. SIMS
Princeton University, USA

MARK F. J. STEEL
University of Warwick, UK

THANASIS STENGOS
University of Guelph, Canada

THOMAS M. STOKER
Massachusetts Institute of Technology, USA

TIMO TERÄSVIRTA
Aarhus University, Denmark

TIMOTHY J. VOGELSANG
Michigan State University, USA

MARK W. WATSON
Princeton University, USA

KENNETH D. WEST
University of Wisconsin-Madison, USA

CHARLES H. WHITEMAN
University of Iowa, USA

TAO ZHA
Federal Reserve Bank of Atlanta, USA

General Preface

All economists of a certain age remember the "little green books". Many own a few. These are the offspring of *The New Palgrave: A Dictionary of Economics*; collections of reprints from *The New Palgrave* that were meant to deliver at least a sense of the *Dictionary* into the hands of those for whom access to the entire four volume, four million word set was inconvenient or difficult. *The New Palgrave Dictionary of Economics, Second Edition* largely resolves the accessibility problem through its online presence. But while the online search facility provides convenient access to specific topics in the now eight volume, six million word *Dictionary of Economics*, no interface has yet been devised that makes browsing from a large online source a pleasurable activity for a rainy afternoon. To our delight, *The New Palgrave*'s publisher shares our view of the joys of dictionary-surfing, and we are thus pleased to present a new series, the "little blue books", to make some part of the *Dictionary* accessible in the hand or lap for teachers, students, and those who want to browse. While the volumes in this series contain only articles that appeared in the 2008 print edition, readers can, of course, refer to the online *Dictionary* and its expanding list of entries.

The selections in these volumes were chosen with several desiderata in mind: to touch on important problems, to emphasize material that may be of more general interest to economics beginners and yet still touch on the analytical core of modern economics, and to balance important theoretical concerns with key empirical debates. The 1987 Eatwell, Milgate and Newman *The New Palgrave: A Dictionary of Economics* was chiefly concerned with economic theory, both the history of its evolution and its contemporary state. The second edition has taken a different approach. While much progress has been made across the board in the 21 years between the first and second editions, it is particularly the flowering of empirical economics which distinguishes the present interval from the 61 year interval between Henry Higgs' *Palgrave's Dictionary of Political Economy* and *The New Palgrave*. It is fair to say that, in the long run, doctrine evolves more slowly than the database of facts, and so some of the selections in these volumes will age more quickly than others. This problem will be solved in the online *Dictionary* through an ongoing process of revisions and updates. While no such solution is available for these volumes, we have tried to choose topics which will give these books utility for some time to come.

Steven N. Durlauf
Lawrence E. Blume

Introduction

This collection of entries covers one of the most important changes in economic methodology between the 1987 and 2008 editions of the *New Palgrave*, namely the nature and role of time series analysis. This explosion has two sources. First, the 1987 edition did not reflect the status of rational expectations as a central feature of modern macroeconomic analysis. As a result, the edition had little on the work of Lars Hansen, Thomas Sargent, and Christopher Sims which has come to dominate current empirical macroeconomics. In contrast, the 2008 edition has much material on topics ranging from the cross-equations restrictions generated by rational expectations, generalized methods of moments approaches to estimation which reflect the interpretation of data interrelationships as first order conditions, and vector autoregression methods that summarize the dymamic properties of data.

Second, time series econometrics experienced a quantum leap in technical sophistication, as exemplified in the work by Robert Engle, John Geweke, Clive Granger, Peter Philips and Peter Robinson. This new work represents a substantial relaxation of the statistical assumptions that had previously been imposed on economic data, as exemplified in the work on ARCH models, unit roots and cointegration and long memory. These new approaches not only have examined the domain of empirical processes which may be subjected to formal analysis but have extended the sorts of questions macroeconomists ask. Examples of this include cointegration, which provides a way of formalizing long run restrictions on comovements across data series that may or may not be consistent with economic theory and unit roots, which changes the way that trend/cycle distinctions are formulated. So, while these entries are among the most mathematically sophisticated of those that appear in the new edition, they are of great relevance to empirical work.

Steven N. Durlauf
Lawrence E. Blume

Contents

aggregation (econometrics)

Aggregation refers to the connection between economic interactions at the micro and the macro levels. The micro level refers to the behaviour of individual economic agents. The macro level refers to the relationships that exist between economy-wide totals, averages or other economic aggregates. For instance, in a study of savings behaviour refers to the process that an individual or household uses to decide how much to save out of current income, whereas the aggregates are total or per-capita savings and income for a national economy or other large group. The econometrics of aggregation refers to modelling with the individual–aggregate connection in mind, creating a framework where information on individual behaviour together with co-movements of aggregates can be used to estimate a consistent econometric model.

In economic applications one encounters many types and levels of aggregation: across goods, across individuals within households, and so on. We focus on micro to macro as outlined above, and our 'individual' will be a single individual or a household, depending on the context. We hope that this ambiguity does not cause confusion.

At a fundamental level, aggregation is about handling detail. No matter what the topic, the microeconomic level involves purposeful individuals who are dramatically different from one another in terms of their needs and opportunities. Aggregation is about how all this detail distils in relationships among economic aggregates. Understanding economic aggregates is essential for understanding economic policy. There is just too much individual detail to conceive of tuning policies to the idiosyncrasies of many individuals.

This detail is referred to as individual heterogeneity, and it is pervasive. This is a fact of empirical evidence and has strong econometric implications. If you ignore or neglect individual heterogeneity, then you can't get an interpretable relationship between economic aggregates. Aggregates reflect a smear of individual responses and shifts in the composition of individuals in the population; without careful attention, the smear is unpredictable and uninterpretable.

Suppose that you observe an increase in aggregate savings, together with an increase in aggregate income and in interest rates. Is the savings increase primarily arising from wealthy people or from those with moderate income? Is the impact of interest rates different between the wealthy and others? Is the response different for the elderly than for the young? Has future income for most people become more risky?

How could we answer these questions? The change in aggregate savings is a mixture of the responses of all the individuals in the population. Can we disentangle it to understand the change at a lower level of detail, like rich versus poor, or young versus old? Can we count on the mixture of responses underlying aggregate savings to be stable? These are questions addressed by aggregation.

Recent progress on aggregation and econometrics has centred on explicit models of individual heterogeneity. It is useful to think of heterogeneity as arising from three broad categories of differences. First, individuals differ in tastes and incomes. Second, individuals differ in the extent to which they participate in markets. Third, individuals differ in the situations of wealth and income risk that they encounter depending on the market environment that exists. Our discussion of recent solutions is organized around these three categories of heterogeneity. For deeper study and detailed citations, see the surveys by Blundell and Stoker (2005), Stoker (1993) and Browning, Hansen and Heckman (1999).

The classical aggregation problem provides a useful backdrop for understanding current solutions. We now review its basic features, as originally established by Gorman (1953) and Theil (1954). Suppose we are studying the consumption of some product by households in a large population over a given time period t. Suppose that the quantity purchased q_{it} is determined by household resources m_{it}, or 'income' for short, as in the formula:

$$q_{it} = \alpha_i + \beta_i m_{it}$$

Here α_i represents a base level consumption, and β_i represents household i's marginal propensity to spend on the product.

For aggregation, we are interested in what, if any, relationship there is between average quantity and average income:

$$q_t = \frac{1}{n_t}\sum_{i=1}^{n_t} q_{it} \quad \text{and} \quad m_t = \frac{1}{n_t}\sum_{i=1}^{n_t} m_{it}$$

where all households have been listed as $i = 1, \ldots, n_t$. Let's focus on one version of this issue, namely, what happens if some new income becomes available to households, either through economic growth or a policy. How will the change in average quantity purchased $\Delta\bar{q}$ be related to the change in average income $\Delta\bar{m}$?

Suppose that household i gets Δm_i in new income. Their change in quantity purchased is the difference between purchases at income $m_{it} + \Delta m_i$ and at income m_{it}, or

$$\Delta q_i = \beta_i \cdot \Delta m_i$$

Now, the average quantity change is $\Delta\bar{q} = \sum_i \Delta q_i / n_t$, so that

$$\Delta\bar{q} = \frac{1}{n_t}\sum_{i=1}^{n_t} \beta_i \cdot \Delta m_i \qquad (1)$$

In general, it seems we need to know a lot about who gets the added income – which i's get large values of Δm_i and which i's get small values of Δm_i. With a transfer policy, any group of households could be targeted for the new income, and their specific set of values of β_i would determine $\Delta\bar{q}$. A full schedule of how much new income goes to each household i as well as how they spend it (that is, Δm_i and β_i), seems like a lot of detail to keep track of, especially if the population is large. Can we ever get by knowing just the change in average income $\Delta\bar{m} = \sum_i \Delta m_i / n_t$?

There are two situations where we can, where a full schedule is not needed:

1. Each household spends in exactly the same way, namely, $\beta_i = \beta$ for all i, so that who gets the new income doesn't affect $\Delta\bar{q}$.
2. The distribution of income transfers is restricted in a convenient way.

Situation 1 is (common) micro linearity, which is termed *exact aggregation*. Another way to understand the structure is to write (1) in the covariance formulation:

$$\Delta\bar{q} = \bar{\beta} \cdot \Delta\bar{m} + \frac{1}{n_t}\sum_{i=1}^{n_t}(\beta_i - \bar{\beta}) \cdot (\Delta m_i - \Delta\bar{m}) \tag{2}$$

where we denote the average spending propensity as $\bar{\beta} = \sum_i \beta_i / n_t$. With exact aggregation there is no variation in β_i, so that $\beta_i = \beta = \bar{\beta}$ and the latter term always vanishes. That is, it doesn't matter who gets the added income because everyone spends the same way. When there is variation in β_i, matters are more complicated unless it can be assured that the new income were always given to households in a way that is uncorrelated with the propensities β_i. 'Uncorrelated transfers' provide an example of a Situation 2, but that is a distribution restriction that is hard to verify with empirical data.

Under uncorrelated transfers, we can also interpret the relationship between $\Delta\bar{q}$ and $\Delta\bar{m}$, that is, the macro propensity is the average propensity $\bar{\beta}$. There are other distributional restrictions that give a constant macro propensity, but a different one from the parameter produced by uncorrelatedness. For instance, suppose that transfers of new income always involved fixed shares of the total amount. That is, household i gets

$$\Delta m_i = s_i \Delta\bar{m} \tag{3}$$

In this case, average purchases are

$$\Delta\bar{q} = \frac{1}{n_t}\sum_{i=1}^{n_t}\beta_i \cdot (s_i \Delta\bar{m}) = \tilde{\beta}_{wtd} \cdot \Delta\bar{m} \tag{4}$$

where $\tilde{\beta}_{wtd}$ is the weighted average $\tilde{\beta}_{wtd} \equiv \sum_i \beta_i s_i / n_t$. This is a simple aggregate relationship, but the coefficient $\tilde{\beta}_{wtd}$ applies only for the distributional scheme (3); it matters who gets what share of the added income. Aside from being a weighted average of $\{\beta_i\}$, there is no reason for $\tilde{\beta}_{wtd}$ to be easily interpretable – for instance, if households with low β_i's have high s_i's, then $\tilde{\beta}_{wtd}$ will be low. If your aim was to estimate the average propensity $\bar{\beta}$, there is no reason to believe that the bias $\tilde{\beta}_{wtd} - \bar{\beta}$ will be small.

Empirical models that take aggregation into account apply structure to individual responses and to allowable distributional shifts. Large populations are modelled, so that compositional changes are represented via probability distributions, and expectations are used instead of averages (for example, mean quantity $E_t(q)$ is modelled instead of the sample average q_t). Individual heterogeneity is the catch-all term for individual differences, and they must be characterized. Distribution restrictions must be applied where heterogeneity is important. For instance, in our example structure on the distribution of new income is required for dealing with the heterogeneity in β_i, but not for the heterogeneity in α_i.

Progress in empirical modelling has come about because of the enhanced availability of micro data over time. The forms of behavioural models in different research areas have been tightly characterized, which is necessary for understanding how to account for aggregation. That is, when individual heterogeneity is characterized empirically, the way is clear to understanding what distributional influences are relevant and must be taken into account. We discuss recent examples of this below.

Some solutions to aggregation problems

Demand models and exact aggregation

It is well known that demand patterns of individual households vary substantially with whether households are rich or poor, and vary with many observable demographic characteristics, such as household (family) size, age of head and ages of children, and so on. As surveyed in Blundell (1988), traditional household demand models relate household commodity expenditures to price levels, total household budget (income) and observable household characteristics. Aggregate demand models relate (economy-wide) aggregate commodity expenditures to price levels and the distribution of income and characteristics in the population. Demand models illustrate exact aggregation, a practical approach for accommodating heterogeneity at the micro and macro levels. These models assume that demand parameter values are the same for all individuals, but explicitly account for observed differences in tastes and income.

For instance, suppose we are studying the demand for food and we are concerned with the difference in demands for households of small size versus large size. We model food purchases for household i as part of static allocation of the budget m_{it} to $j=1,\ldots,J$ expenditure categories, where food is given by $j=1$, and price levels at time t are given by $P_t=(p_{1t},\ldots,p_{Jt})$. Small families are indicated by $z_{it}=0$ and large families by $z_{it}=1$.

Expenditure patterns are typically best fit in budget share form. For instance, a translog model of the food share takes the form

$$w_{1it}=\frac{p_{1t}q_{1it}}{m_{it}}=\frac{1}{D(p_t)}\left[\alpha_1+\sum_{i=1}^{J}\beta_{1j}\ln p_{jt}+\beta_m\ln m_{it}+\beta_z z_{it}\right] \tag{5}$$

where $D(p_t)=1+\sum_{i=1}^{J}\beta_j\ln p_{jt}$. The parameters ($\alpha_1$ and all β's) are the same across households, and the price levels (p_{jt}'s) are the same for all households but vary with t. Individual heterogeneity is represented by the budget m_{it} and the family size indicator z_{it}. We have omitted an additive disturbance for simplicity, which would represent another source of heterogeneity. The important thing for aggregation is that model (5) is intrinsically linear in the individual heterogeneity. That is, we can write

$$w_{1it} = b_1(p_t) + b_m(p_t) \cdot \ln m_{it} + b_z(p_t) \cdot z_{it} \tag{6}$$

The aggregate share of food in the population is the mean of food expenditures divided by mean budget, or

$$W_{1t} = \frac{E_t(m_{it}\, w_{1it})}{E_t(m_{it})} = b_1(p_t) + b_m(p_t) \cdot \frac{E_t(m_{it}\ \ln\ m_{it})}{E_t(m_{it})} + b_z(p_t) \cdot \frac{E_t(m_{it} z_{it})}{E_t(m_{it})} \tag{7}$$

The aggregate share depends on prices, the parameters (α_1 and all β's) and two statistics of the joint distribution of m_{it} and z_{it}. The first,

$$S_{mt} = \frac{E_t(m_{it} \ln m_{it})}{E_t(m_{it})} \tag{8}$$

is an entropy term that captures the size distribution of budgets, and the second

$$S_{zt} = \frac{E_t(m_{it} z_{it})}{E_t(m_{it})} \tag{9}$$

is the percentage of total expenditure accounted for by households with $z_{it} = 1$, that is, large families.

The expressions (6) and (7) illustrate *exact aggregation* models. Heterogeneity in tastes and budgets (incomes) are represented in an intrinsically linear way. For aggregate demand, all one needs to know about the joint distribution of budgets m_{ti} and household types z_{it} is a few statistics; here S_{mt} and S_{zt}.

The obvious similarity between the individual model (6) and the aggregate model (7) raises a further question. How much bias is introduced by just fitting the individual model with aggregate data, that is, putting $E_t(m_{it})$ and $E_t(z_{it})$ in place of m_{it} and z_{it}, respectively? This can be judged by the use of *aggregation factors*. Define the factors π_{mt} and π_{zt} as

$$\pi_{mt} = \frac{S_{mt}}{\ln E_t(m_{it})} \quad \text{and} \quad \pi_{zt} = \frac{S_{zt}}{E_t(z_{it})}$$

so that the aggregate share is

$$W_{1t} = \frac{E_t(m_{it}\, w_{1it})}{E_t(m_{it})} = b_1(p_t) + b_m(p_t) \cdot \pi_{mt} \cdot \ln E_t(m_{it}) + b_z(p_t) \cdot \pi_{zt} \cdot E_t(z_{it})$$

One can learn about the nature of aggregation bias by studying the factors π_{mt} and π_{zt}. If they are both roughly equal to 1 over time, then no bias would be introduced by fitting the individual model with aggregate data. If they are roughly constant but not equal to 1, then constant biases are introduced. If the factors are time varying, more complicated bias would result. In this way, with exact aggregation models, aggregation factors can depict the extent of aggregation bias.

The current state of the art in demand analysis uses models in exact aggregation form. The income (budget) structure of shares is adequately represented as quadratic in $\ln m_{it}$, as long as many demographic differences are included in the analysis. This means that aggregate demand depends explicitly on many statistics of the income-demographic distribution, and it is possible to gauge the nature and sources of aggregation bias using factors as we have outlined. See Banks, Blundell and Lewbel (1997) for an example of demand modelling of British expenditure data, including the computation of various aggregation factors.

Exact aggregation modelling arises naturally in situations where linear models have been found to provide adequate explanations of empirical data patterns. This is not always the case, as many applications require models that are intrinsically nonlinear. We now discuss an example of this kind where economic decisions are discrete.

Market participation and wages

Market participation is often a discrete decision. Labourers decide whether to work or not, firms decide whether to enter a market or exit a market. There is no 'partial' participation in many circumstances, and changes are along the extensive margin. This raises a number of interesting issues for aggregation.

We discuss these issues using a simple model of labour participation and wages. We consider two basic questions. First, how is the fraction of working (participating) individuals affected by the distribution of factors that determine whether each individual chooses to work? Second, what is the structure of average wages, given that wages are observed only for individuals who choose to work? The latter question is of interest for interpreting wage movements: if average wages go up, is that because (*a*) most individual wages went up or (*b*) low-wage individuals become unemployed, or leave work? These two reasons give rise to quite different views of the change in economic welfare associated with an increase in average wages.

The standard empirical model for individual wages expresses log wage as a linear function of time effects, schooling and demographic (cohort) effects. Here we begin with

$$\ln w_{it} = r(t) + \beta \cdot S_{it} + \varepsilon_{it} \tag{10}$$

where $r(t)$ represents a linear trend or other time effects, S_{it} is the level of training or schooling attained by individual i at time t, and ε_{it} are all other idiosyncratic factors. This setting is consistent with a simple skill price model, where $w_{ti} = R_t H_{it}$ with skill price $R_t = e^{r(t)}$ and skill (human capital) level $H_{it} = e^{\beta S_{it} + \varepsilon_{it}}$. We take eq. (10) to apply to all individuals, with the wage representing the available or offered wage, and β the return to schooling. However, we observe that wage only for individuals who choose to work.

We assume that individuals decide whether to work by first forming a reservation wage

$$\ln w_{it}^{*} = s^{*}(t) + \alpha \ln B_{it} + \beta^{*} \cdot S_{it} + \zeta_{it}$$

where $s(t)$ represents time effects, B_{it} is the income or benefits available when individual i is out of work at time t, S_{it} is schooling as before, and ζ_{it} are all other individual factors. Individual i will work at time t if their offered wage is as big as their reservation wage, or $w_{it} \geq w_{it}^{*}$. We denote this by the participation indicator I_{it}, where $I_{it} = 1$ if i works and $I_{it} = 0$ if i doesn't work. This model of participation can be summarized as

$$\begin{aligned} I_{it} &= 1\,[w_{it} \geq w_{it}^{*}] = 1[\ln w_{it} - \ln w_{it}^{*} \geq 0] \\ &= 1\,[s(t) - \alpha \ln B_{it} + \gamma \cdot S_{it} + \nu_{it} \geq 0] \end{aligned} \tag{11}$$

where $s(t) \equiv r(t) - s^{*}(t)$, $\gamma \equiv \beta - \beta^{*}$ and $\nu_{It} \equiv \varepsilon_{it} - \zeta_{it}$.

If the idiosyncratic terms ε_{it}, ν_{it} are stochastic errors with zero means (conditional on B_{it},S_{it}) and constant variances, then (10) and (11) is a standard selection model. That is, if we observe a sample of wages from working individuals, they will follow (10) subject to the proviso that $I_{it} = 1$. This can be accommodated in estimation by assuming that ε_{it}, ν_{it} have a joint normal distribution. That implies that the log wage regression of the form (10) can be corrected by adding a standard selection term as

$$\ln w_{it} = r(t) + \beta \cdot S_{it} + \frac{\sigma_{\varepsilon\nu}}{\sigma_{\nu}} \lambda \left[\frac{s(t) - \alpha \ln B_{it} + \gamma\, S_{it}}{\sigma_{\nu}} \right] + \eta_{t}. \tag{12}$$

Here, σ_{ν} is the standard deviation of ν and $\sigma_{\varepsilon\nu}$ is the covariance between ε and ν. $\lambda(\cdot) = \phi(\cdot)/\Phi(\cdot)$ is the 'Mills ratio', where ϕ and Φ are the standard normal p.d.f. and c.d.f respectively. This equation is properly specified for a sample of working individuals – that is, we have $E(\eta_t | S_{it}, B_{it}, I_{it} = 1) = 0$. For a given levels of benefits and schooling, eq. (11) gives the probability of participating in work as

$$E_t[I_{it} | B_{it}, S_{it}] = \Phi \left[\frac{s(t) - \alpha \ln B_{it} + \gamma \cdot S_{it}}{\sigma_{\nu}} \right] \tag{13}$$

where $\Phi[\cdot]$ is the normal c.d.f.

For studying average wages, the working population is all individuals with $I_{it} = 1$. The fraction of workers participating is therefore the (unconditional) probability that $\alpha \ln B_{it} - \gamma \cdot S_{it} - \nu_{it} \leq s(t)$. This probability is the expectation of I_{it} in (11), an intrinsically nonlinear function in observed heterogeneity B_{it} and S_{it} and unobserved heterogeneity ν_{it}, so we need some explicit distribution assumptions. In particular, assume that the participation index $\alpha \ln B_{it} - \gamma \cdot S_{it} - \nu_{it}$ is normally distributed with mean $\mu_t = \alpha\, E_t(\ln B_{it}) - \gamma E_t(S_{it})$ and variance

$$\sigma_t^2 = \alpha^2 Var_t(\ln B_{it}) + \beta^2 Var_t(S_{it}) - 2a\beta \cdot Cov_t(\ln B_{it}, S_{it}) + \sigma_{\nu}^2. \tag{14}$$

Now we can derive the labour participation rate (or one minus the unemployment rate) as

$$E_t[I_{it}] = \Phi\left[\frac{s(t) - \alpha E_t(\ln B_{it}) + \gamma E_t(S_{it})}{\sigma_t}\right] \tag{15}$$

where again $\Phi[\cdot]$ is the normal c.d.f. This formula relates the participation rate to average out-of-work benefits $E_t(\ln B_{it})$ and average training $E_t(S_{it})$, as well as their variances and covariances through σ_t. The specific relation depends on the distributional assumption adopted; (15) relies on normality of the participation index in the population.

For wages, a similar analysis applies. Log wages are a linear function (10) applicable to the full population. However, for participating individuals, the intrinsically nonlinear selection term is introduced, so that we need explicit distributional assumptions. Now suppose that log wage $\ln w_{it}$ and the participation index $\alpha\ln B_{it} - \gamma \cdot S_{it} - v_{it}$ are joint normally distribution. It is not hard to derive the expression for average log wages of working individuals

$$\begin{aligned} E_t[\ln w_{it}|I_{it} = 1] = r(t) + \beta \cdot E_t(S_{it}|I = 1) \\ + \frac{\sigma_{\varepsilon v}}{\sigma_t}\lambda\left[\frac{s(t) - \alpha E_t(\ln B_{it}) + \gamma E_t(S_{it})}{\sigma_t}\right] \cdot \end{aligned} \tag{16}$$

This is an interesting expression, which relates average log wage to average training of the workers as well as to the factors that determine participation.

However, we are not interested in average log wages, but rather average wages $E_t(w_{it})$. The normality structure we have assumed is enough to derive a formulation of average wages, although it is a little complex to reproduce in full here. In brief, Blundell, Reed and Stoker (2003) show that the average wages of working individuals $E\ [w_{it}|I_{it} = 1]$ can be written as

$$\ln\ E[w_{it}|I_{it} = 1] = r(t) + \beta \cdot E_t(S_{it}) + \Omega_t + \Psi_t \tag{17}$$

where Ω_t, Ψ_t are correction terms that arise as follows. Ω_t corrects for the difference between the log of an average and the average of a log, as

$$\Omega_t \equiv \ln E_t(w_{it}) - E_t(\ln w_{it}) + \Omega_t.$$

Ψ_t corrects for participation, as

$$\Psi_t \equiv \ln E[w_{it}|I_{it} = 1] - \ln E_t\ (w_{it}).$$

Recall our original question, about whether an increase in average wages is due to an increase in individual wages or to increased unemployment of low-wage workers. That is captured in (17). That is, Ψ_t gives the participation effect, and the other terms capture changes in average wage $E_t(w_{it})$ when all are participating. As such, this