On state dependence:
a case for two-handed economists

Stephen P. Jenkins (LSE)

Email: s.jenkins@lse.ac.uk

International Workshop on 'Dynamics of Low Wages, Low Pay, and Transfer Receipt'

IAB, Nuremberg, 15 November 2013

This research draws on work undertaken with Lorenzo Cappellari (Università Cattolica Milano), and Herwig Immervoll and Sebastian Königs (Social Policy Division, ELSA, OECD), but they bear no responsibility for any of the views expressed here. Part-supported by the UK Economic and Social Research Council (grant RES-535-25-0090) and the University of Essex through core-funding of the Research Centre on Micro-Social Change.

State dependence: working definition

Binary outcome describing status at time t ($y_{it} = 1$ or 0)

- E.g. low paid/not low paid; poor/not poor; receiving SA benefit/not receiving; employed/not employed; unemployed/not unemployed; …

Aggregate (raw) state dependence (*ASD*)

- Probability of being in current state is greater for those previously in that state than those who were not
  - Difference in raw transition probabilities

$$\Pr(y_{it} = 1 | y_{it-1} = 1) - \Pr(y_{it} = 1 | y_{it-1} = 0)$$

Genuine state dependence (*GSD*)

- Probability of being in current state is causally related to whether previously in that state, i.e. controlling for associations arising from *heterogeneity*, observed or unobserved (or other endogenous selections)
  - Model-based assessment: averaging over predicted probabilities

$$\text{APE} = \text{Average}[\ \Pr(y_{it} = 1 | y_{it-1} = 1, \text{heterogeneity}) - \Pr(y_{it} = 1 | y_{it-1} = 0, \text{heterogeneity})\ ]$$

Example 1: low pay in Britain

Stewart & Swaffield, *Economica*, 1999

Men and women, BHPS waves 1–5

Using e.g. low pay threshold 3 (= ⅔ hourly median),

- ASD = 73 ppt
- GSD = 56 ppt (76% of GSD)

TABLE 6

DEPENDENCE OF PROBABILITY OF BEING LOW PAID IN YEAR t ON LOW-PAY STATUS IN YEAR $t-1$

	P(LP in t | LP in $t-1$)			P(LP or not E in t | LP in $t-1$)		
Threshold	1	2	3	1	2	3
Raw aggregate probabilities of low paid in year t, given						
1. Low paid at $t-1$	0·648	0·736	0·801	0·706	0·773	0·827
2. Higher paid at $t-1$	0·042	0·061	0·074	ASD	0·121	0·128
3. Difference	0·606	0·675	0·727	[illegible]	0·652	0·699
Model predicted probabilities: estimates of P(LP in year t | LP in year t − 1)						
Probit model						
4. Average over low paid at $t-1$ sample	0·648	0·737	0·801	0·706	0·773	0·827
5. Average over higher paid at $t-1$ sample	0·428	0·539	0·648	0·520	0·594	0·683
6. Difference	0·220	0·198	0·154	0·186	0·179	0·144
7. state dependence effect	0·386 (64%)	0·477 (71%)	0·574 (79%)	0·412 (69%)	0·473 (73%)	0·555 (79%)
Endogenous selection model						
8. Average over low paid at $t-1$ sample	0·648	0·737	0·801	0·706	0·773	0·827
9. Average over higher paid at $t-1$ sample	0·393	0·513	0·630	0·484	0·569	0·663
10. Difference	0·255	0·224	0·171	0·222	0·204	0·165
11. state dependence effect	0·351 (58%)	0·452 (67%)	0·556 (76%)	GSD	0·448 (69%)	0·534 (76%)

Notes:
(a) Calculations based on estimated models given in Table 5.
(b) LP denotes low paid and E employment.

Example 2: unemployment, British men

Arulampalam, Booth, and Taylor, *OEP*, 2000

British men, BHPS waves 1–5

- Aged 25+: ASD ≈ 50 ppt, GSD ≈ 20 ppt (40%)
- Aged < 25: ASD ≈ 50 ppt, GSD ≈ 10 ppt (20%)

Table 4 Raw data probabilities and predicted unemployment probabilities from base model—variant (i) (1991–1995)

	Wave 1 (initial conditions)		Wave 2		Wave 3		Wave 4		Wave 5	
	Aged under 25 [1]	Aged 25+ [2]	Aged under 25 [3]	Aged 25+ [4]	Aged under 25 [5]	Aged 25+ [6]	Aged under 25 [7]	Aged 25+ [8]	Aged under 25 [9]	Aged 25+ [10]
Raw Data Probabilities										
(1) Average over age groups	0.0894	0.0763								
(2) Unemployed at $t-1$			0.5345	0.5313	0.5833	0.5479	0.6667	0.5417	0.3500	0.5472
(3) Employed at $t-1$			0.0299	0.0397	0.0150	0.0358	0.0144	0.0292	0.0078	0.0177
(4) (2) − (3)		**ASD**	0.5046	0.4916	0.5683	0.5121	0.6523	0.5125	0.3422	0.5295
Predicted probabilities holding characteristics constant										
(5) Unconditional	0.0126	0.0168								
(6) Unemployed at $t-1$			0.1683	0.2698	0.1619	0.2431	0.1388	0.2159	0.1038	0.1529
(7) Employed at $t-1$			0.0515	0.0440	0.0456	0.0320	0.0357	0.0252	0.0246	0.0136
(8) State dependence (6) − (7)			0.1168	0.2258	0.1163	0.2111	0.1031	0.1907	0.0792	0.1393
As a percentage of (4)		**GSD**	23.15	45.93	20.46	41.22	15.81	37.21	23.14	26.31

Interpretation and policy implications

- The greater is GSD, the greater the policy pay-off to generic measures related to the state of interest rather than to measures related to individual characteristics

Identification of the extent of true state dependence among men of working age is more than just an academic exercise. There is considerable diversity of government policies towards the unemployed in market economies throughout the world.[2] These policies rest on very different assumptions about the extent to which government intervention can alter the equilibrium, or so-called natural rate, of unemployment. If there is no state dependence in unemployment incidence at the micro level, then short run policies to reduce unemployment (such as job creation schemes and wage subsidies) will have no effect on the equilibrium aggregate unemployment rate.

But if there is true state dependence, then policies reducing short run unemployment incidence will have longer run effects by reducing the natural rate of unemployment. The prevention of the initial unemployment experience becomes an important policy objective, perhaps indicating the need to focus on education and training initiatives. Thus it is crucial to know whether there is genuine state dependence, or if instead it is individual heterogeneity that causes individuals to be repeatedly observed in unemployment. The increasing flexibility of the labour market

of reservation wages for those who experience low pay. As well as shedding light on labour market functioning, the distinction between heterogeneity and GSD is relevant for policy: in the pure heterogeneity case, policies targeted according to the covariates of persistence can reduce entrapment into low-paid jobs, while in the pure state dependence case, universal measures such as minimum wages are the more appropriate tool. The situation in the real world is somewhere in the middle between these two polar cases, and the analysis of low pay mobility is helpful for understanding where.

or constraints caused by previous working behavior. For the purpose of policy evaluation, it is critical to determine the relative contributions of state dependence and individual characteristics to the observed persistence in labor supply behavior. Indeed, if labor supply choices are driven entirely by observed or unobserved individual characteristics, then the effect of a policy intervention, such as a wage subsidy or an in-work benefit, will cease the moment the policy is withdrawn. In contrast, if past labor market outcomes exert a causal effect on current labor supply behavior, then the policy intervention will affect labor market outcomes beyond the duration of the policy.

The implications for policy design of those two scenarios are very different. If welfare receipt *per se* increases the probability of future benefit receipt, policies that prevent entry or facilitate exit from social assistance can induce a lasting reduction in welfare dependency. If, in contrast, high welfare persistence is a consequence of the recipients' individual characteristics, policies that encourage exits from social assistance will have little impact as long as these barriers to self-sufficiency continue to exist. In terms of the concepts introduced by Heckman (1981a; 1981b), the question whether observed persistence in social assistance is 'structural' or whether it results from unobserved individual heterogeneity and is hence purely 'spurious' is therefore of considerable relevance for policy makers. ¶

One-handed economists?

“Give me a one-handed economist! All my economists say, On the one hand on the other.” Harry S. Truman

- A plea for clear policy guidance

Two-handed economists!

- This presentation argues that our approach to SD has been too one-handed, albeit in a different sort of way
- On the one hand, we've made advances in the econometric modelling of dynamic discrete choice models incorporating SD and heterogeneity
  - More and better models; more and more applications
- But, on the other hand, the concrete policy implications of this research are arguably of little use to policy-makers
  - We need to know much more about the underlying mechanisms that generate SD
  - And thereby become rather more two-handed

Outline of rest of presentation

1. Review of econometric models for estimating GSD

2. The case for more about mechanisms, and hence better policy-relevance

Modelling approaches

1. Dynamic random effects probit (DREP) models
   - From Heckman (1981), Hyslop (*E'metrica* 1999), etc. etc., onwards
   - E.g. Arulampalam et al. (*OEP* 2000), as earlier
   - E.g. papers at 2013 OECD/IZA Conference on Safety Nets …
   - The most commonly-used, so will get the most discussion
2. Endogenous switching models
   - E.g. Stewart and Swaffield (*Economica* 1999), as earlier;
   - Cappellari and Jenkins (2004, 2008*a*, *b*)
3. Linear probability models
   - Used to supplement DREP models (cf. Stewart, *JAE* 2007)
4. Multi-state models
   - Increasing number of applications
   - Prowse (*JBES* 2012) model of FTE, PTE, NE is the most statistically sophisticated so far

DREP model: basic specification

Latent propensity of binary outcome for person i in each year t of the sequence of T_i years linearly related to lagged binary outcome, observed characteristics, (time-invariant) unobserved individual characteristics, and random error

$$p^*_{it} = \lambda y_{it-1} + \beta' Z_{it-1} + \gamma_i + \zeta_{it}; \quad t = 2, \ldots, T_i$$

- GSD monotonically increasing in λ, assumed to be common to all i, and fixed over time
- RE (random intercept) model
  - But dynamic FE model problematic
- Simple dynamic structure: FO Markov (1 lag); no persistence in idiosyncratic shocks

DREP model: issues

- Initial conditions problem (Heckman 1981)
  - Correlation between lagged outcome and error term leads to bias: state in which first observed in panel (y_{i1}) related to same unobservables
- Largely resolved nowadays
  - Heckman (1981): model the initial condition jointly (approximate model); requires instruments and special software
  - Orme (1997, 2001): two-step procedure, easy to implement
  - Wooldridge (2005): condition on the set of explanatory variables and y_{i1}, modelling γ_i as a function of time-averaged (or all observed) and y_{i1} (and possible their interaction). Easy to implement
  - Each estimator provides similar results for long panels (Heckman best for short panel); unbalanced panel not a big deal
    - Arulampalam & Stewart (OBES 2009), Akay (*JE'metrics 2009*), Cappellari & Jenkins (OECD WP, 2008)

DREP model issues (continued)

- Attrition
  - Not considered much (and note that models developed assuming balanced panels)
  - Possible to use longitudinal weights if assume MAR
    - Wooldridge (2002), and Biewen (2009)
- Normality of errors (probit model)
  - Not a big deal: logit or mass point heterogeneity seems to deliver similar results
- Heterogeneity in SD under-explored
  - Unlikely to be constant across individuals or calendar time (policy effects?)
  - Variation with observables: interactions with (elements of) Z, or separate models for subgroups (constrained by cell sizes)
  - Variation with unobservables: e.g. random coefficient on y_{it-1}, possibly correlated with random intercept (Stewart 2007; Prowse 2012). Non-trivial to estimate

DREP model issues (continued)

- Dynamics specification
  - Estimate of SD is conditional on how persistence elsewhere in the model is modelled and, moreover, why only one lag and not more?
  - More general specifications for error term such as AR(1): Hyslop (1999), Chay and Hyslop (2000), Stewart (1997), Prowse (2012), … or even unrestricted correlation matrix
  - E.g. Andrén & Andrén (2013) on SA benefit receipt: λ biased upwards in basic model relative to more general specifications
  - Lags of second as well as first order: e.g. Stewart (1997), Prowse (2012), Andrén & Andrén (2013)
  - Use of Wooldridge estimator facilitates fitting of these models
  - More complicated specifications require long panels to be fitted well – identification of 'persistence' requires long window of observation!

DREP model issues (continued)

- Strict exogeneity assumption regarding observed predictors (Z)
  - Probability of outcome at t, conditional on Z and initial condition, is unrelated to past or future values of $Z \Rightarrow$ can express density of data in form for estimation
  - Rules out feedback effects; ignoring them biases estimates of GSD (Biewen, *JAE* 2009, poverty in Germany)
    - Household composition predicts current poverty, but poverty now may affect future household composition; similarly with employment status
    - Model ('endogenise') the problematic predictors: relate them to past outcome realisations (y) … but can't have fully simultaneous model (Heckman)
    - Identified using instruments (plausible?) and functional form; special software
    - GSD measured by APE of lagged poverty status is 61 ppt in basic model and only 31 ppt in model with feedback effects
  - General lessons?
    - For econometricians
    - For policy-makers

Endogenous switching models

Latent propensity of binary outcome for person *i* in each year *t*:

$$p_{it}^{*} = [(y_{it-1})\gamma_1' + (1-y_{it-1})\gamma_2']Z_{it-1} + \varepsilon_{it};\ \varepsilon_{it} = \tau_i + \zeta_{it}$$

- Impact of *Z* differs ('switches') according to whether $y_{it-1} = 1$ or $y_{it-1} = 0$
- Auxiliary equation to model initial conditions à la Heckman; equation errors are bivariate normal
- Can add additional equations to model other 'endogenous selections', e.g. attrition from sample
  - E.g. Cappellari and Jenkins (*JAE* 2004) on GB poverty, and their later papers on low pay (2008*a*, *b*)
  - Buddlemeyer & Verick (*Econ Rec* 2008) on AUS poverty
  - Aretz & Gürtzgen (IZA DP, 2012): IAB admin data for DE, low pay with time-varying GSD and employment retention
  - NB C & J software in Stata for MSL estimation with *K*-variate normal

Endogenous switching models: issues

- GSD not a single number: heterogeneity built into approach
- Fitted by pooling data on pairs of years over panel
  - Requires plausible instruments for each of the initial conditions and other selection equations (otherwise identification by functional form)
  - Need to account for correlations between error terms across obs (repeated obs across individuals): use cluster-robust SE estimator, where clusters are defined by e.g. wave 1 household identifier
  - Inefficient estimates relative to those derived by modelling full *sequence* of outcomes: doesn't fully exploit panel
    - Cf. Ribar's (*SEJ* 2005) more complicated model using sequences
  - But circumvents the strict exogeneity issue (Biewen 2009)!
  - Can extend to allow ordered categories for lagged outcome rather than binary, modelling IC as ordered probit (C & J 2004)
    - But complicated …

Dynamic Linear Probability models

- Feasible approach to allowing for more general assumptions about error structure using DPD models (Anderson-Hsiao, Arellano-Bond, Blundell-Bond)
  - ‘fixed effect’ idea: unobserved effects differenced out
  - E.g. Stewart (*JAE* 2007): leads to differences in estimated APE compared to DREP APE, but small

- Potential issue: LP models versus binary outcome models in general
- Potential issue: commonly-found sensitivity of estimates from DPD models

Multi-state dynamic models

Richer models: from binary discrete outcome to categorical outcome; from lagged binary state to lagged multiple states

- SD now potentially encompasses cross-state effects, and one can test whether dependence same for different lagged states
- Dynamic multinomial logit models with unobs het and Wooldridge (or Heckman) IC estimator, e.g.
  - NE|E-Formal|E-Informal (Gong et al., *EDCC* 2004); Inactivity|E|welfare (Wunder and Riphahn, *OEP* 2013); NE|Low Pay|High Pay (Fok et al., Melbourne WP 2013; Mosthaf et al. IZA DP 2009); NE|E-PT|E-FT (Prowse, *JBES* 2012)
  - Wunder-Riphahn: SOEP, WG natives and immigrants; 'no strong evidence' for GSD in welfare-welfare and similar to inactivity-welfare
  - Prowse: BHPS 18 waves, women; 2 lags, random coefficients and intercepts, and AR(1) errors; 'omission of random coefficients and autocorrelation biases estimates of state dependencies'; 'part-time employment does not appear to entail lower labour-market attachment compared with full-time employment' (despite being worse jobs)

Multi-state dynamic models (continued)

- MNL models now most common (software available for all but most complicated), but they're not the only ones:
- Dynamic bivariate RE probit model (generalised DREP)
  - Stewart (*JAE* 2007; WP 2005): 'Low pay as a conduit to repeat unemployment' (similar effects on current employment chances of past no pay and past low pay)
  - outcomes are low pay (high pay), and unemployment; lagged values of low pay and unemployment (BHPS waves 1–6, men and women)
  - See also Knabe and Plum (*Labour* 2013), SOEP
    - but lagged low pay observed only if not unemployed last period
- Endogenous switching model, with multiple endogeneous selections
  - Cappellari and Jenkins (*RLE* 2008): outcomes are men's employment, low pay (and high pay), with multiple selections (BHPS waves 1–10, men)
    - Predicted *E* status depends on past *E* status (and past *LP* status if E_{t-1}=1); predicted *LP* status depends on past *E* status (and past *LP* status if E_{t-1}=1)
    - More evidence for low-paid work affecting employment chances à la past unemployment

Two generic (data) issues

- Is it state dependence or some other sort of dependence?
- Heckman & Borjas (1980): 4 types of dependence (in discussion of whether past unemployment causes current unemployment)
  1. Markovian dependence (as discussed)
  2. Occurrence dependence (# previous spells)
  3. Duration dependence (length of time in state since entry)
  4. Lagged duration dependence (length of previous spells)
- Ability to investigate effects of each depends on nature of longitudinal data available
  - Panel point-in-time; interval-censored; genuine continuous (dated) transitions
    - Cf. Bhuller, Brich, and Königs (2013): estimates from FO Markov fitted to monthly data on benefit receipt, and implications for annual derived and then compared with estimates from model fitted to annual data. Find that more aggregated models tend to over-estimate degree of GSD

State dependence or duration dependence?

- The 'continuing spell' issue: studies based on household panels with y_{it} referring to status at date of annual interview tend to find a significant fraction of respondents in state had been in that state since at least the previous interview (according to between-interview retrospective histories)
- So, are DREP model estimates of λ simply picking up effect of continuing spell rather than genuine SD?
- Substantially lower GSD estimates if omit observations with continuing spells, according to BHPS examples:
  - Arulampalam, Booth, and Taylor (*OEP* 2000)
  - Stewart (*JAE* 2007), APE for lagged UN: 1.5 rather than 3.5 ppt
  - Cappellari and Jenkins (OECD WP 2008): APE for lagged SA ⅓ as big

The definition of the (discrete) state at *t*, *t*–1

- All our estimates, including of GSD, are conditional on having 'valid' definitions of the states of interest
- To what extent is discretization artificial?
  - 'Poor', when poverty lines are largely arbitrary
  - 'Low paid', when low pay cut-offs are largely arbitrary
  - 'Employment' more clear cut? But part-time versus full-time?
    - 30 hours convention versus (UK) 16 hours definition for benefits purposes
  - 'Social assistance benefit receipt', more clear cut
    - Except that benefit systems change over time, so hard to use consistent definition
- Does it make sense to think of 'state dependence' in terms of (lagged) discrete/categorical measures?
- Empirical checks satisfactory?
  - Repeat analysis using different thresholds
  - Differential effects with ordered category lag var (C & J 2004)
  - Model outcome and its lag using continuous measures?

What have we learnt?

On the one hand, …

- For virtually every outcome considered, researchers find evidence of significant non-zero GSD
  - the size of which is, however, substantially less than ASD
  - GSD magnitude varies by outcome, country, calendar time period
- Notable substantial advances in statistical techniques
- Facilitated in part by growing availability of longitudinal data
- Multi-state model provide nuanced perspectives on these results, but essentially the same headline about GSD's existence
- There remains plenty for one-handed modellers to work on, including addressing assumptions about unobserved heterogeneity (cf. Prowse) and feedback effects (cf. Biewen)

What have we learnt?

On the other hand, …

- Rather less consideration given to the two generic but rather fundamental issues
  - NB Heckman & Borjas (1980) strategies for distinguishing different types of dependence require even richer data
- Only superficial investigation of the behavioural mechanism(s) underlying GSD
  - the literature to date has provided rather brief and general explanations of 'why past receipt matters'; and usually these explanations could be applied to either duration dependence or Markovian state dependence!
- Analysis of mechanisms complicated by overlaps between the various outcomes discussed so far …
  - Unemployed, low paid, poor, benefit recipient

More on GSD mechanisms

Social assistance (SA) benefit example:

- Overlaps: whether a working-age individual receives SA benefits is the result of an administrative decision about eligibility where, by definition, eligibility depends on the income of the claimant's family or household
- Low income arises from low pay or especially unemployment
- Also, SA benefit levels in most countries would give SA recipients an income lower than or similar to the official poverty line
- So, to what extent does state dependence in SA receipt reflect dependence in these other domains (with associated policy implications), and are there factors associated with past SA receipt alone?
  - See Contini and Negri (*ESR* 2007): simulation model to illustrate how SA dependence may be zero and arise entirely from SD in e.g. poverty or unemployment
- GSD in (un)employment an obvious source of GSD in SA receipt

GSD in (un)employment: mechanisms

[See inter alia Heckman and Borjas (1980), Arulampalam, Booth, and Taylor (2000), and Stewart (2007), and references therein]

1. Human capital depreciation: having no job can mean that a worker’s existing training and educational skills and experience may lose their labour market value and opportunities to update them on the job are unavailable and, in turn, these effects increase the chance of future unemployment

GSD in (un)employment: mechanisms

2. If employers screen potential employees on the basis of their unemployment histories (over and above other characteristics such as their education and skills)
   - Past unemployment provides a cheap signal to employers regarding low labour productivity, with adverse consequences for the individuals concerned.
   - Important evidence about such signalling in US field experiment by Kroft, Lange and Notowidigdo *(QJE* 2013)
     - Posting fake CVs to real job applications in 100 US cities and tracking call-backs; design uses variation in applicants' unemployment spell lengths; stronger negative effect of longer spell when local labour market tight
   - This aspect may be hard to distinguish from state dependence in low pay if there is significant cycling of workers between low pay and unemployment, and low-waged jobs – in addition to unemployment – do not maintain or enhance workers' human capital, or are used as a screening device by employers

GSD in (un)employment: mechanisms

3. Preferences may change with the experience of unemployment:
   - ‘individuals in unemployment may lower their reservation wage with the passage of time, and accept poorer quality jobs that are more likely to be destroyed, and for this reason may be more likely to experience unemployment in the future’ (Arulampalam, Booth, and Taylor 2000: 26)
4. If effective intensive job search is reliant on financial expenditure in addition to investments of time, and unemployment leads to a substantial running down of financial assets and savings, then past unemployment may reduce the chances of re-employment

- Arguably, the driver in the third and fourth cases might be poverty rather than unemployment (Contini and Negri 2007: 25)
  - I.e. it is a lack of financial means more generally that leads to poor quality jobs being taken or prevents job search activity

GSD in (un)employment: mechanisms

- ‘Expectancy’ and ‘cultural’ theories of dependence (Bane and Ellwood 1994)
  - Another explanation for a change in preferences underlying a lowering of reservation wages over time
  - ‘Expectancy’ theory: adverse impacts of unemployment on individuals’ confidence and feelings of self-control, motivation, and self-esteem which then have adverse effects on job-finding
    - There might also be a deleterious feedback loop from lack of job finding to psychological factors
    - Mosthaf (2013): changes in ‘self-efficacy’ as a source of GSD in employment, using German PASS data … ordered logit for self-efficacy and binary logit for employment, with lags of each in each equation, estimated jointly; … but ‘the impact of employment on self-efficacy and vice versa is close to 0’

GSD in (un)employment: mechanisms

- ‘Cultural theory’ ≈ peer or neighbourhood effects
  - the idea that social groups can have powerful norms, which individuals within the group would find it difficult to deviate from
  - Being unemployed among many unemployed people may be normal; getting a regular job may be abnormal. So, social pressures of various kinds may lead people to change their attitudes change if they become unemployed
  - These arguments have greatest plausibility in the context of ghettos (Bane and Ellwood 1994).
  - But even then it is difficult to claim even in principle that it is unemployment that is the principal driver, since ghettos are locations with an intense concentration of disadvantage of various kinds, including poverty and low paid work as well as unemployment and benefit receipt

GSD in (un)employment: mechanisms

- It may be a neighbourhood culture of poverty or of benefit receipt, rather than or as well as unemployment which changes preferences
- A different type of social group effect of unemployment: adverse impact on the size and nature of the circle of social contacts that help people find out about jobs or to get them
  - These effects are more likely the result of unemployment if the relevant contacts are also in work; otherwise arguably similar effects might arise from a lack of income to afford to socialise
    - Bane and Ellwood (1994) were discussing the reason for ‘welfare dependence’ in 1980s USA when the principal welfare (SA) benefit was AFDC, and the recipients were mostly black lone mothers

GSD in SA benefit receipt: mechanisms

Are there any 'pure' SA GSD mechanisms, i.e. other than via (un)employment?

- Most plausible argument for a 'pure' SA effect is likely to be that SA receipt stigmatises its recipients in ways that unemployment or poverty do not
  - E.g. having psychosocial effects on recipients, or potential employers treating histories of SA receipt differently from, say, receipt of contributory unemployment insurance benefits
  - Otherwise difficult to make a strong case for pure SA receipt effects in general?
    - May well differ from country to country, and depend inter alia on …
    - the particular procedures involved in claiming SA benefits (how demeaning are they?)
    - the availability of different sorts of benefits (assistance and insurance), and
    - the prevalence and concentration of receipt (stigma is likely to be less the more that one's peers are also in receipt)

GSD mechanisms: conclusions

- We know remarkably little about these mechanisms
- Remarkably little research on this compared to the econometric research on the other hand?
- Surely, we need to know more about mechanisms in each domain/output in order to inform policy-makers?
- There are different implications depending on which SD mechanism is the most important

So,

- We need fewer studies simply documenting the magnitude of GSD
- We need more studies investigating mechanisms
  - i.e. more like Kroft, Lange and Notowidigdo (*QJE* 2013) and Mosthaf (2013)
  - Hard(er) to do, but more innovative and pay-off higher

LSE

In sum, we need to be more two-handed!

Selected references

Akay, A. (2012). ‘Finite-sample comparison of alternative methods for estimating dynamic panel data models’, *Journal of Applied Econometrics*, 17, 1189–1204.

Andrén, T. and Andrén, D. (2013). ‘Never give up? The persistence of welfare participation in Sweden, *IZA Journal of European Labor Studies*, 2:1.

Aretz, B. and Gürtzgen, N. (2012). ‘What explains the decline in mobility in the German low-wage sector?’, Discussion Paper 7046, IZA, Bonn.

Arulampalam, W., Booth, A. L., and Taylor, M. P. (2000). ‘Unemployment persistence’, *Oxford Economic Papers*, 52, 24–40.

Arulampalam, W. and Stewart, M. B. (2009). ‘Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators’, *Oxford Bulletin of Economics and Statistics*, 71(5), 659–681.

Bane, M.J. and Ellwood, D.T. (1986). ‘Slipping into and out of poverty: the dynamics of spells’, *Journal of Human Resources*. 21(1), 1–23.

Bane, M. J. and Ellwood, D. T. (1994). *Welfare Realities: From Rhetoric to Reform*. Cambridge MA: Harvard University Press.

Bhuller, M. and Königs, S. (2011). ‚The dynamics of social assistance receipt in Norway’, unpublished paper, University of Oxford and Statistics Norway

Bhuller, M., Brinch, C. and Königs, S. (2013). ‘Time aggregation and the analysis of welfare persistence’, unpublished paper, University of Oxford and Statistics Norway.

Biewen, M. (2009). ‘Measuring state dependence in individual poverty histories when there is feedback to employment status and household composition’, *Journal of Applied Econometrics*, 24: 1095–1116.

Buddelmeyer, H. and Verick, S. ‘Understanding the drivers of poverty dynamics in Australian households’, *Economic Record* 84 (266), 310–321.

Cappellari, L. (2007). ‘Earnings mobility among Italian low-paid workers’, *Journal of Population Economics*, 20, 465–482.

Cappellari, L. and Jenkins, S. P. (2004). ‘Modelling low income transitions’, *Journal of Applied Econometrics*, 19(5), 593–610.

Cappellari, L. and Jenkins, S. P. (2008a). ‘The dynamics of social assistance receipt: measurement and modelling issues, with an application to Britain’, *Social, Employment and Migration Working Paper 67*. Paris: OECD.

Cappellari, L. and Jenkins, S. P. (2008b). ‘Estimating low pay transition probabilities accounting for endogenous selection mechanisms’, *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 57(2), 165–186.

Cappellari, L. and Jenkins, S. P. (2008c).‘Transitions between low pay and unemployment’, Chapter 8, in S. Polachek and K. Tatsiramos (eds), *Research in Labor Economics, Volume 28*, Elsevier, Amsterdam, 57–79.

Cappellari, L. and Jenkins, S. P. (2013). ‘The dynamics of social assistance benefit receipt in Britain’. Revised version of IZA Discussion Paper 4457, presented at IZA/OECD/World Bank Conference on ‘Safety Nets and Benefit Dependence: Evidence and Policy Implications’, OECD, Paris, 21–22 May 2013.

Chay, K. Y. and Hyslop, D. (2000). ‘Identification and estimation of dynamic binary response panel data models: empirical evidence using alternative approaches’, unpublished Paper, University of California at Berkeley.

Contini, D. and Negri, N. (2007). ‘Would declining exit rates from welfare provide evidence of welfare dependence in homogeneous environments?’ *European Sociological Review*, 23 (1), 21–33.

Ellwood, D. T. (1982). ‘Teenage unemployment: permanent scars or temporary blemishes’, in: R. B. Freeman and D. A. Wise (eds), *The Youth Labor Problem: Its Nature, Causes and Consequences*. Chicago: University of Chicago Press.

Finnie, R. and Pavlic, D. (2013). ‘The dynamics of social assistance receipt in Canada’. unpublished working paper, University of Ottawa

Fok, Y. K., Scutella, R., and Wilkins, R. ‘The low-pay, no-pay cycle: are there systematic differences across demographic groups?, Melbourne Institute Working Paper 32/13, University of Melbourne.

Gong, X,. van Soest, A., Villagomez, E. (2004).. ’Mobility in the urban labor market: a panel data analysis for Mexico’, *Econ omic Development and Cultural Change*, 53 (1), 1–36.

Hansen, H.T. (2009). ‘The dynamics of social assistance recipiency: empirical evidence from Norway’, *European Sociological Review*, 25(2), 215–231.

Hansen, J. and Lofstrom, M. (2006). ‘Immigrant-native differences in welfare participation: the role of entry and exit rates’, IZA Discussion Paper No. 2261.

Hansen, J., Lofstrom, M., and Zhang, X. (2006). ‘State dependence in Canadian welfare participation’, IZA Discussion Paper No. 2266.

Heckman, J. (1981a). ‘Heterogeneity and state dependence’, in Rosen (ed.), *Studies in Labor Markets*. University of Chicago Press .

Heckman, J. J. (1981b). ‘The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process’, in: C. F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press.

Heckman, J. J. (1981c). ‘Statistical models for discrete panel data’, in: C. F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press.

Heckman, J. J. and Borjas, G. J. (1980). ‘Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence’, *Economica*, 47, 247–283.

Knabe, A. and Plum, A. (2013). ‘Low wage jobs – springboard to high-paid ones?’, *Labour*, 27 (3), 310–330 .

Königs, S. (2012). ‘The dynamics of social assistance benefit receipt in Luxembourg – a descriptive analysis’, INET Working Paper, University of Oxford. http://www.emod.ox.ac.uk/sites/emod.ox.ac.uk/files/Konigs_2012.pdf

Königs, S. (2013a). ‘The dynamics of social assistance benefit receipt in Germany’, *OECD Social, Employment and Migration Working Papers*, 136 (forthcoming). OECD, Paris

Königs, S. (2013b). ‘The dynamics of social assistance benefit receipt in the Netherlands’, *OECD Social, Employment and Migration Working Papers*, forthcoming. OECD, Paris

Kroft, K., Lange, F., and Notowidigdo, M. J. (2013). ‘Duration dependence and labor market conditions: evidence from a field experiment’, *Quarterly Journal of Economics*, online advanced access

Mosthaf, A. (2013). ‘Change in self-efficacy as a source of state dependence in labor market dynamics’, unpublished paper, University of Mainz, Mainz

Mosthaf, A., Schank, T., and Schnabel, C. (2009). ‘Low wage employment versus unemployment. Which one provides better prospects for women?’, Discussion Paper 4611. Bonn: IZA.

Mundlak, Y. (1978). ‘On the pooling of time series and cross section data’, *Econometrica*, 46(1), 69–85.

Orme, C.D. (2001). ‘Two-step inference in dynamic non-linear panel data models’, unpublished paper, University of Manchester. http://personalpages.manchester.ac.uk/staff/chris.orme/documents/ Research%20Papers/initcondlast/pdf

Prowse, V. (2012). ‘Modeling employment dynamics with state dependence and unobserved heterogeneity’, *Journal of Business and Economic Statistics*, 30 (3), 411–431.

Ribar, D. C. (2005). ‘Transitions from welfare and the employment prospects of low-skill workers’, *Southern Economic Journal*, 71 (3), 514–533.

Stewart, M. B. (2007). ‘The interrelated dynamics of unemployment and low pay’, *Journal of Applied Econometrics*, 22(3), 511–531.

Stewart M. B, and Swaffield, J. K. (1999). ‘Low pay dynamics and transition probabilities’, *Economica*, 66, 23–42.

Wooldridge, J. M. (2002). ‘Inverse probability weighted M-stratification’, *Portuguese Economic Journal*, 1, 117–139.

Wooldridge, J. M. (2005). ‘Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity’, *Journal of Applied Econometrics*, 20(1), 39–54.

Wunder, C. and Riphahn, R. (2013). ‘ The dynamics of welfare entry and exit among natives and immigrants’, *Oxford Economic Papers*, online first.