A Multiple Indicators and Multiple Causes (MIMIC) Model of Immigrant Settlement Success

Paper Presented to the XXVI International Population Conference of the IUSSP

October 2, 2009

Laurence H. Lester B.Ec (Hons) M.Ec, PhD

Ph: +61 8 8201 2002
Fax: +61 8 8276 9060
Email: laurence.lester@flinders.edu.au

NILS
National Institute of Labour Studies
Flinders University, Adelaide, Australia

MODELLING IMMIGRANT SUCCESSFUL SETTLEMENT[1]

This paper presents models of immigrant successful settlement by application of a linked Multiple Indicators and Multiple Causes (MIMIC) model, a special case of a longitudinal structural equation model (SEM), in which the influences of formative indicators on unobservable latent variables are assessed through their impact on the reflective indicators. As no well-documented explicit model of successful settlement exists, the devised statistical models provide the first comprehensive assessment in a framework that simultaneous assesses multiple dimensions of the immigrant settlement process.

Models of immigrant successful settlement are constructed for the two cohorts of the Longitudinal Surveys of Immigrants to Australia (LSIA). The LSIA data were an important initiative of Department of Immigration, Multicultural and Indigenous Affairs (DIMIA)[2] and are considered to be "…world class surveys of recent migrants…" (Richardson *et al*. 2002, p.5) which are a rich and comprehensive source of information about immigrants to Australia that "…are particularly well-suited to addressing the dynamics of settlement…" (Cobb-Clark 2001, p.468).

Reflective Indicators of Successful Settlement

The four reflective indicators of successful settlement (SucSet) are provided in Table 1.

Table 1: Successful Settlement–Reflective Effects-Indicators (LSIA)

Measure	Variable Name
Level of satisfaction with life in Australia	*LifeOk*
Mental Health (GHQ-12)	*GHQ*
Decision to immigrate was right	*RightMig*
Encourage others to migrate to Australia	*Encore*

Notes: (1) See later discussion for indicator scaling issues.

Formative Indicators: The Causes of Successful Settlement

The set of formative (causal) indicators of SucSet are listed in Table 2 below. The labour market is represented by the index of labour market success (LMSI)—see Lester (2006) for details. Thus, the labour market impact on SucSet is mediate through the LMSI by inclusion of that index as a formative indicator for immigrants who are labour force participants (Graff & Schmidt 1985).

[1] This paper is an edited extract of chapter 5 of my 2008 (unpublished) PhD thesis. I thank the ARC and DIMIA for funding.

[2] The first survey was collected on behalf of the Department of Immigration and Multicultural Affairs (DIMA), which subsequently became DIMIA.

All indicators and variables (which have order) are coded so that larger values are better or preferred (e.g. Spons takes values of 1 if sponsored and 0 if not). In addition, ordinal data are standardised (mean zero, unit standard deviation) based on the underlying unobserved continuous variable, and continuous data are similarly rescaled for comparable units.[3]

Two variables are treated differently from time-variant variables that are available in all waves. The wave 2 and wave 3 variables BetOff (comparing the household's current income with their income in the previous wave) and BetHome (a comparison of current housing standard with that in the previous wave) are not available in wave 1. In principle, they enter models as formative indicators in wave 2 (and 3 for LSIA1).

[3] In the LISREL software, ordinal data are standardised (zero mean, unit standard deviation) prior to calculation of the correlation matrix on which analysis is based. To ensure comparability of results, non-ordinal data are treated similarly so that distinctly differential units are not responsible for results.

Table 2: Causes of Successful Settlement (LSIA)

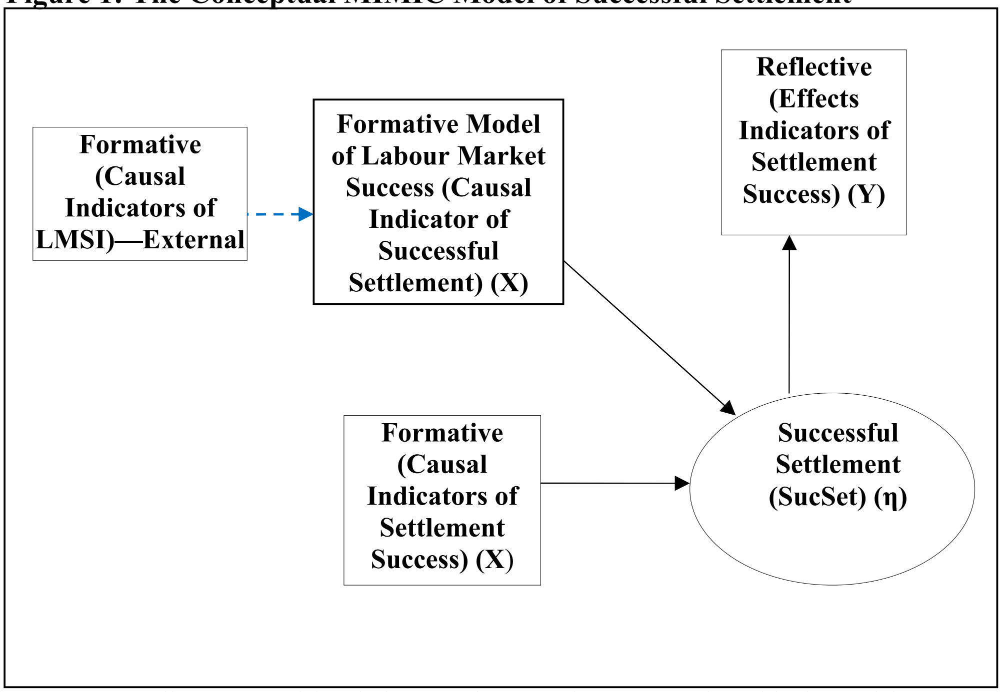
Measure	Variable Name	Settlement Domain
CAUSAL FACTORS COMMON TO ALL RESIDENTS		
Age	*Age*	NA
Better off financially now	*BetOff*	Financial Well-being
Education	*Educat*	NA
Gender	*Gender*	NA
Home ownership vs. not owner (mortgage, rent, other)	*OwnHome*	Financial Well-being
Housing standard improved	*BetHome*	Financial Well-being
Marital status	*Marstat*	Non-specific—May contribute to social participation
Number of Adults in the household	*NumAdult*	Non-specific: Social participation?
Number of children in the household	*NumChild*	Non-specific: Social participation?
Physical Health (Number of visits to a Doctor)	*DrVisit*	Health
Relative Income	*Relinc*	Financial Well-being
Wealth	*Wealth*	Financial Well-being
CAUSAL FACTORS SPECIFIC TO IMMIGRANTS		
Choice of Australia was influenced by employment, or economic conditions	*CameEco*	Non-Specific
Choice of Australia was to join relatives or to marry	*CameFam*	Non-Specific: May contribute to social participation
Cultural similarity (replaces Country of birth or origin) and/or English-speaking developed country (U.S.A., U.K., Ireland, Canada)	*PDI* and/or *EngBack*	Non-Specific: May contribute to social participation
English Language ability index (see note 1)	*ELAI*	Social Participation
Sponsored	*Spons*	Social Participation
Time in Australia since arrival	*TimeOz*	Social Participation
LABOUR MARKET SPECIFIC (Incorporated in the LMSI)—See Lester (2006)		
Income (from wage & salary jobs & from all sources – per hour and in levels) (Employed & Unemployed)	*W&SInc* & *IncAll*	Labour Market
Labour market status	*Nowlfs*	Labour Market
Like versus dislike job (Employed)	*JobSat*	Labour Market
Looking for a replacement for main job (Unemployed)	*Lookjob*	Labour Market
Occupational status (Converted to the ANU4) (Employed)	*OccStat*	Labour Market
Perceived difficulty in finding a job (Unemployed)	*DiffJob*	Labour Market
Receiving an unemployment benefit (Unemployed)	*Umpbfit*	Labour Market
Receiving assistance in finding work (Unemployed)	*HelpJob*	Labour Market

Notes: (1) The ELAI is an index formed using LSIA questions asked about the immigrant's ability to speak, read, and write English.

The Conceptual Model

The conceptual model forming the basis of analysis of successful settlement (SucSet) is demonstrated as the stylised path diagram of a MIMIC model of SucSet in Figure 1.

Figure 1: The Conceptual MIMIC Model of Successful Settlement



Notes: (1) Details suppressed for clarity. (2) The LMSI is constructed outside the model—represented by the dashed-arrow. (3) Formative (**X**) indicators may be time-variant or time-invariant, Reflective indicators (**Y**) are time-variant. (4) SucSet is the latent endogenous (η) variable representing successful settlement.

A MIMIC model has a formative structural model, and a reflective measurement model. Following convention, for formative indicators, single-headed arrows lead from indicators to the latent construct SucSet, The dashed-arrow for the formative LMSI represents construction of the LMSI outside the MIMIC model. For reflective indicators arrows lead from SucSet to the indicators.

The Formative Measurement Model

In the formative model, the is hypothesised is that SucSet is influenced by, for example Gender (male or female) and Person (whether the immigrant was a primary applicant (PA) or migrating unit spouse or partner (MU)). Formative indicators are assumed to be correlated and to be measured without error—which in most cases is uncontentious.

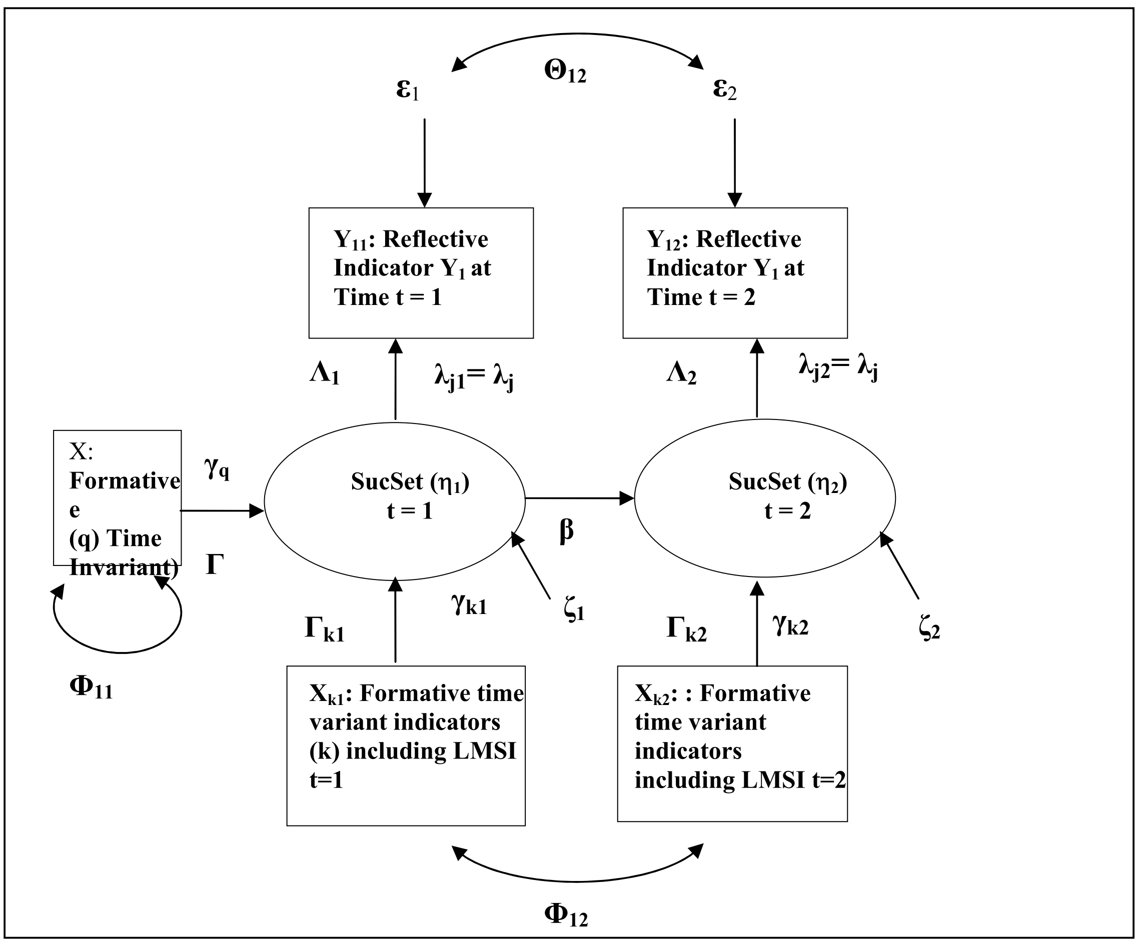
The Reflective (Factor Analytical) Model

Reflective indicators' errors (ε) are correlated across time and are assumed to contain measurement error.

The MIMIC Model of SucSet

A two-period (i.e. for LSIA2) path diagram of a MIMIC model of SucSet combining reflective and formative components is given in Figure 2 below. Thus, the MIMIC model permits simultaneous estimation of the measurement model and the incorporation of causal variables in the structural model for the latent variable SucSet: SucSet is linearly determined (apart from random errors, ζ) by formative indicators or variables—and SucSet determines the observed reflective indicators (apart from random errors, ε).

Figure 2: Linked MIMIC Model



Notes: (1) Multiple time-invariant causal indicators represented by **X**—with the vector of γ path coefficients (Γ). (2) $\mathbf{X}_{k1}$ and $\mathbf{X}_{k2}$ are time-variant causal indicators at time 1 and time 2—with vectors of γ_k path coefficients (Γ_k). (3) Double-headed arrows represent correlated errors: Θ_{12} represents the matrix of θ correlations between reflective indicator errors (ε) at time 1 and 2; Φ_{11} represent the matrix of ϕ cross-sectional correlations for formative indicators at time (wave) one; Φ_{12} represents the matrix of ϕ longitudinal correlations for formative indicators. (4) Path coefficients for reflective time-invariant indicators are equal at each point in time (e.g. (in Λ) $\lambda_{11} = \lambda_{12} = \lambda_1$). (5) ε and ζ are measurement errors.

Figure 2 above demonstrates an important model issue relating to longitudinal models. Factor loadings (path coefficients) for time-variant reflective indicators are constrained to be equal across time (e.g. for reflective indicator the vector (Λ) of *j* coefficients at time 1 and time 2 are equal: $\lambda_{j1}= \lambda_{j2}= \lambda_j$). This ensures that changes in SucSet are due to changes in indicators not factor loadings (Jöreskog 2004).

For ordinal reflective indicators, it is also necessary to ensure that indicators are measured on the same scale at each wave. This is accomplished by setting equal thresholds for the underlying (unobserved) variables—represented by the observed ordinal reflective indicators—prior to generating the correlations[4] on which analysis is based (Jöreskog and Sörbom 2002; Jöreskog 2004; Brown 2006). Note that this is in addition to specifying a unit loading for one reflective indicator to set the scale of the latent variable—see below.

Several exogenous time-invariant formative variables are specified as influencing SucSet at wave 1 (e.g. Gender). These time-invariant variables are treated as states or history, they influence SucSet at time periods beyond the first through their impact on the initial level of SucSet—through the β coefficients (see Figure 2 above) (De Leeuw *et al.*; Markowitz 2001; Kim and Rojewski 2002; Hellgren and Sverke 2003). In addition, note that for LSIA1, allowing the β coefficient to vary (β_{12}, the structural coefficient between SucSet1 and SucSet2, is not constrained to equal β_{23}) allows statistical assessment of the stability of the relationship between SucSet at various points in time—see below.

The LSIA Data

Sub-Group Models

Groups of particular interest in the analysis of successful settlement here are economic immigrants (subject to a points test), non-economic immigrants, and those who are not labour force participants (NLF). Table 3 gives data for the interaction between labour force participation, economic, and non-economic immigrants in the LSIA.

[4] Polychoric, tetrachoric, polyserial, or biserial as appropriate.

Table 3 Labour Force Participants and Economic Immigrant (LSIA)

LSIA1	Economic	%	Non-Economic	%
Labour force participants	997	20.5	2598	53.4
Not in labour force	18	0.4	1253	25.7
Total	**1015**	**20.9**	**3852**	**79.1**
LSIA2	**Economic**	**%**	**Non-Economic**	**%**
Labour force participants	1416	40.0	1062	30.0
Not in labour force	71	2.0	990	28.0
Total	**1487**	**42.0**	**2051**	**58.0**

Notes: (1) Data are weighted. (2) Totals may not add due to rounding.

Given the small number of economic-NLF immigrants, this group cannot be analysed separately. Instead, three groups are constructed, economic immigrants (who are all labour force participants), non-economic immigrants who are labour force participants (henceforth referred to as non-economic), and NLF: disaggregation results in three exclusive groups as shown in Table 4. As this Table shows, the proportion of economic immigrants in LSIA2 was about twice that in LSIA1—a consequence of changes to immigrant selection policy and access to welfare benefits (see Lester 2008).

Table 4: Immigrant Groups (LSIA)

	LSIA1	%	LSIA2	%
Economic Immigrants in Labour Force	997	20.5	1416	40.0
Non-Economic Immigrants in Labour Force	2598	53.4	1062	30.0
Non-Labour Force Participants (NLF)	1271	26.1	1060	30.0
Total	**4867**	**100**	**3538**	**100**

Notes: (1) Data are weighted. (2) Totals may not add due to rounding.

SucSet of economic immigrants is expected to be influenced by labour market outcomes: more generally, for labour market participants (either economic or non-economic immigrants), labour market influences on SucSet are examined by including the LMSI as a causal indicator in MIMIC models to follow.

Model Assessment

Table 5 below provides the model goodness-of-fit statistics that are applicable to MIMIC models to follow.

Table 5: Model Fit Assessment and Test Statistics

Test Statistic	Purpose	Acceptance Criteria
Root Mean-Square Error of Approximation (RMSEA)	Absolute fit (0 is perfect fit, < 0.01 is outstanding)	< 0.05 close < 0.08 good < 0.10 reasonable
Standardised Root Mean-square Residual (SRMR)	Absolute fit	< 0.10 favourable < 0.05 good
Goodness-of-Fit (GFI) Adjusted Goodness-of-Fit (AGFI)	Absolute fit (range 0 no fit, 1 perfect fit)	> 0.90 good fit
Comparative Fit Index (CFI)	Incremental fit (range 0 to 1)	> 0.90 good fit
Parsimony-based GFI (PGFI) & Parsimony-based Normed Fit Index (PNFI)	Incremental, parsimony adjusted, fit (range 0 to 1)	No defined level
Akaike Information Criterion (AIC) and Consistent Akaike Information Criterion (CAIC)	Comparative model fit (no upper limit, 0 perfect fit)	No defined level
Expected Value of the Cross-validation index (ECVI)	Comparative model fit (no upper limit, 0 perfect fit)	No defined level
Chi-squared (χ^2)	Comparative model fit (see note 2)	No defined level

Notes: (1) The Chi-squared statistic is not a reliable goodness-of-fit indicator in large samples, but it is useful to assess the relative fit of various models (Brown 2006)—the Chi-squared statistic is the Satorra-Bentler Scaled Chi-squared, which takes non-normality of input data into account).

Models are based on analysis of the correlation matrix—goodness-of-fit is an assessment of how well the derived model replicates the observed correlation matrix—and, as there is no single goodness-of-fit measure, it is practice in applied work to report several appropriate statistics.

Model Derivation

A two-step modelling approach is used to construct MIMIC models. First, the reflective measurement model of SucSet is considered using exploratory and confirmatory factor analysis. When an appropriate structure is suggested for the measurement model, the full MIMIC model is considered—that is, the structural model and formative models are added to the measurement model.

Data Screening

The data from the LSIA used to model SucSet are predominantly ordinal (several are dichotomous), and they generally have fewer than seven categories. Analysis of ordinal data requires special techniques (i.e. they should not be treated as continuous), and the data are more likely to result in model estimation problems. The LSIA data were not necessarily collected for the sophisticated modelling undertaken in this paper, and so deficiencies must be seen in context—they are the cost of access to data that provides a unique opportunity to examine the course of immigrant settlement.

Table 6 provides the correlations for the observed reflective indicators of SucSet at each wave of the data: LSIA1 data are above the diagonal and LSIA2 below the diagonal—variable suffix indicates the LSIA wave (with no wave 3 for LSIA2). Correlation matrices for the sub-samples of economic, non-economic, and non-labour force participants (on which models are based) differ, but not to such an extent that the full sample misrepresents the underlying relationships.

Table 6: Reflective Indicator Correlations All Immigrants (LSIA)

Indicator	1	2	3	4	5	6	7	8	9	10	11	12
1 Encore1	1	0.59	-	0.14	0.07	-	0.36	0.29	-	0.55	0.40	-
2 Encore2	0.51	1	-	0.07	0.11	-	0.23	0.29	-	0.43	0.54	-
3 Encore3	0.49	0.55	1	-	-	-	-	-	-	-	-	-
4 GHQ1	0.17	0.13	0.09	1	0.43	-	0.45	0.26	-	0.43	0.31	-
5 GHQ2	0.08	0.10	0.12	0.42	1	-	0.21	0.33	-	0.15	0.35	-
6 GHQ3	0.03	0.04	0.15	0.33	0.40	1	-	-	-	-	-	-
7 LifeOk1	0.37	0.24	0.19	0.41	0.23	0.13	1	0.54	-	0.67	0.44	-
8 LifeOk2	0.26	0.32	0.26	0.32	0.33	0.21	0.46	1	-	0.42	0.69	-
9 LifeOk3	0.19	0.22	0.32	0.23	0.25	0.39	0.34	0.52	1	-	-	-
10 RightMig1	0.52	0.28	0.24	0.45	0.17	0.13	0.62	0.40	0.27	1	0.65	-
11 RightMig2	0.36	0.43	0.33	0.26	0.30	0.19	0.46	0.65	0.50	0.55	1	-
12 RightMig3	0.26	0.27	0.50	0.24	0.25	0.34	0.33	0.53	0.67	0.51	0.68	1

Notes: (1) LSIA2 correlations are above the diagonal, LSIA1 below the diagonal. (2) Correlations are polychoric, tetrachoric, polyserial, or biserial as appropriate.

Model Identification

MIMIC models to follow are identified. As the models are not complex (in terms of latent variables) the "counting rule" (a necessary but not sufficient condition) suggests identification—and LISREL software confirms identification (models solve, and identification problems are not reported). More specifically, as there are a minimum of three statistically significant reflective indicators for SucSet, the SEM model is identified.

Structural Equation Model (Longitudinal Data Model)

Table 7 (LSIA1) and Table 8 (LSIA2) below provide model estimates and goodness-of-fit statistics for the longitudinal MIMIC (hybrid SEM) of SucSet—based on Figure 2 above (since loadings are constrained to be equal across time only one set per cohort need be given).

Individual models for all immigrants, economic immigrants, non-economic immigrants who are labour force participants, and NLF immigrants are examined.

Table 7: Longitudinal Structural Equation Model of Successful Settlement (LSIA1)

LSIA1	All Immigrants (4-Indicator)	Economic Immigrant in Labour Force	Non-Economic Immigrants in Labour Force	Not Labour Force Participants
Structural Model (β) and *t*-Statistic				
SucSet1 → SucSet2	0.700	0.733	0.663	0.781
t-statistic	28.880	20.772	20.453	15.745
SucSet2 → SucSet3	0.743	0.770	0.713	0.837
t-statistic	32.119	17.775	22.209	15.981
Measurement Model (Path Coefficients, λ, and *t*-Statistics)				
Encore	0.573	0.687	0.482	0.702
t-statistic	19.993	18.952	14.286	11.045
GHQ	0.543	0.524	0.523	0.599
t-statistic	27.155	18.745	18.771	13.769
LifeOk	1.0	1.0	1.0	1.0
t-statistic	n.a.	n.a.	n.a.	n.a.
RightMig	1.071	1.118	0.966	1.280
t-statistic	27.223	22.451	19.263	12.902
Model Goodness-of-Fit Statistics and Details				
RMSEA	0.063	0.080	0.065	0.067
SRMR	0.047	0.058	0.049	0.067
GFI	0.992	0.987	0.991	0.986
AGFI	0.988	0.981	0.987	0.979
CFI	0.976	0.967	0.973	0.974
Chi-squared (df)	1054.1 (52)	383.0 (52)	619.1 (52)	344.1 (52)
N	4867	997	2598	1271
AIC	1130.1	459.0	695.1	420.1
CAIC	1414.8	683.4	955.8	653.7
ECVI	0.232	0.461	0.268	0.331
Reliability SucSet1→2	0.542	0.573	0.479	0.717
Reliability SucSet2→3	0.500	0.554	0.456	0.596
Variance SucSet1	0.641	0.650	0.715	0.476
Variance SucSet2	0.580	0.609	0.656	0.405
Variance SucSet3	0.640	0.651	0.731	0.476

Notes: (1) n.a. (not applicable) indicates a *t*-statistic (or standard error) is not available for the "fixed" reference variable. (2) Reliability (i.e. the squared multiple correlations, SMC) is, e.g. the proportion of variance of SucSet2 explained by SucSet1. (3) A dash (-) represents an excluded indicator. (4) Data are weighted. (5) Estimation method is DWLS. (7) Sample size: 4867.

Table 8: Longitudinal Structural Equation Model of Successful Settlement (LSIA2)

LSIA2	All Immigrants (4-Indicator)	Economic Immigrant in Labour Force	Non Economic Immigrants in Labour Force	Not Labour Force Participants
Structural Model (β) and *t*-Statistic				
SucSet1 → SucSet2	0.617	0.684	0.549	0.664
t-statistic	23.864	19.417	13.507	14.131
Measurement Model (Path Coefficients, λ, and t-Statistics)				
Encore	0.639	0.839	0.494	0.712
t-statistic	16.863	15.609	10.992	8.305
GHQ	0.575	0.723	0.557	0.439
t-statistic	22.431	18.796	13.435	10.633
LifeOk	1.0	1.0	1.0	1.0
t-statistic	n.a.	n.a.	n.a.	n.a.
RightMig	1.214	1.462	0.912	1.351
t-statistic	18.873	20.697	12.300	12.010
Model Goodness-of-Fit Statistics and Details				
RMSEA	0.047	0.060	0.071	0.063
SRMR	0.058	0.047	0.075	0.087
GFI	0.992	0.995	0.989	0.977
AGFI	0.986	0.992	0.981	0.962
CFI	0.99	0.986	0.975	0.982
χ^2 (df)	186.7 (21)	132.9 (22)	131.8 (21)	113.2 (22)
N	3538	1416	1062	1060
AIC	232.7	176.9	178.1	157.2
AIC Null	16735.9	8133.4	4516.1	5005.1
CAIC	397.6	314.5	315.3	288.4
CAIC Null	16793.2	8183.5	4563.9	5052.8
ECVI	0.066	0.125	0.168	0.148
ECVI Null	4.732	5.748	4.257	4.726
Reliability SucSet1→2	0.431	0.490	0.348	0.553
Variance SucSet1	0.617	0.484	0.848	0.554
Variance SucSet2	0.546	0.462	0.736	0.441

Notes: (1) n.a. (not applicable) indicates a *t*-statistic (or standard error) is not available for the "fixed" reference variable. (2) A dash (-) represents an excluded indicator. (3) Data are weighted. (4) Sample size: 3538. (5) Estimation method is DWLS. (6) For economic immigrants it is necessary to fix the error variance of RightMig (to a small positive value) to ensure model convergence with non-negative error variance (an "improper solution" for this indicator—see Byrne (1998) for general examples, or Warren *et al*. (2002 for a specific example. Since the resulting model is good in other respects, this can be treated as a data issue not a model misspecification (Brown 2006).

All models in Table 7 and Table 8 proved at least a "good" fit to the data: for both cohorts for all models the RMSEA statistic is less than 0.08 (and in several cases, the 95% confidence interval for the RMSEA includes 0.05 indicating a "close" fit). Other absolute fit statistics are above/below the cut-off points for a good fit. For example, for LSIA2 for all immigrants for the

4-indicator model: RMSEA = 0.047 < 0.05, SRMR = 0.058 < 0.10, GFI = 0.992 > 0.90; AGFI = 0.986 > 0.90, CFI = 0.990 > 0.90), and model fit statistics for comparison with the null model are above the cut-off point (e.g. for LSIA2, AIC = 232.69 < null 16735.9, CAIC = 397.63 < null 16793.2, and ECVI = 0.066 < null 4.73).

Factor loadings for reflective indicators are statistically significant (in the group models, *t*-statistics in LSIA1 range from 11.045 to 22.451 and in LSIA2 from 9.429 to 20.697, i.e. significant at the 0.001% level or better).

The coefficients relating SucSet across waves is also strongly statistically significant (i.e. in the group models the lowest *t*-statistics is 13.494). Moreover, SucSet at later periods can be predicted from the value at previous periods with some degree of accuracy: for example for NLF immigrants in LSIA2 (4-indicator model), the reliability (SMC) for SucSet is 0.553—approximately 55 per cent of the variation in SucSet at C2W2 can be explained by SucSet at C2W1. For all models for sub-groups (economic, non-economic, and NLF), the proportion of SucSet explained by the previous period SucSet ranges from about 35 per cent for non-economic immigrants in LSIA2, to about 72 per cent for NLF immigrants in LSIA1 between waves 1 and 2. Relatively low variance explained suggest formative variables play a greater role, and previous levels of SucSet a lesser role, in predicting SucSet—addressed in the MIMIC models to follow. The ability of previous levels of SucSet to predict later levels of SucSet is consistent with the view that subjective well-being tends to revert to a set-point, or exhibits homeostasis. Thus, later levels of SucSet (which itself measures a broad form of subjective well-being), are partially predictable from current levels.

All groups of immigrants show a statistically significant reduction in the variance of SucSet between wave 1 and 2, but LSIA1 immigrants show an increase in variance between wave 2 and wave 3. Thus, during the first 18 months in Australia, immigrants become more homogeneous with respect to SucSet, but in LSIA1 they tend to become less so as more time passes. Whether this is true for all immigrants to Australia in all periods is beyond the ability of the data to predict (i.e. there is no wave 3 for LSIA2).

The models for the three sub-groups suggest that there are material differences between the groups—in particular, factor loadings (λ) differ and the 95 per cent confidence interval for the estimates do not all coincide.

Consistent with the hypothesis that SucSet can be measured using latent variable models, the longitudinal SEM models discussed above demonstrate that unobserved multi-dimensional SucSet can readily be represented by a number of reflective indicators. Thus, a longitudinal SEM of successful settlement is supported by the data and hence successful settlement can be assessed in a factor analytical model based on observable indicators across time periods. The longitudinal SEM models show the relationship between SucSet across waves appears to be partly dependent on whether immigrants are economic immigrants and whether they are in the labour force. The similarities and differences between groups are explored in the MIMIC models to follow.

MIMIC Model Specification

Having previously established a successful longitudinal SEM for SucSet (incorporating the measurement and structural models), the second stage of estimation is the single-factor, 4-indicator, longitudinal SEM with the inclusion of the formative model—that is the MIMIC model.

MIMIC Modelling Strategy

In comparison with the econometric literature, there is almost no discussion in the literature relating to MIMIC models regarding a modelling strategy—beyond the advice, followed above, that the SEM measurement model precedes the MIMIC model, and a preference for parsimony (Kline 2005). As there are few practical examples of linked (longitudinal) MIMIC models there is also little guidance in the applications literature. Since the arguments for the (top-down) general-to-specific method in the econometrics literature are well-developed, and as this approach results in an econometrically derived parsimonious model (in which irrelevant variables are removed to increase validity of the estimates and of the model assessment statistics) the method is adopted for MIMIC models here. Given that the "causal" part of the MIMIC model is analogous to multiple regression analysis (Kline 2006), the application of the general-to-specific method is appropriate.[5] Finally, the less complex the model, the less information that needs to be collected to examine settlement outcomes for immigrants beyond the LSIA data.

Given that the general-to-specific approach is appropriate, the selection of the cut-off point for removing variables from the model warrants consideration. As the models to be examined are exploratory, prudence suggests the balance between removing variables (parsimony versus omitted variable) and inclusion of irrelevant variables (over-fitting) be relaxed and the usual 5%

[5] Fleishman *et al*. (2002) use a general-to-specific method (referred to as "backward elimination"), and Chung *et al*. (2005) exclude non-significant causal variables from their MIMIC model.

significance level for variable exclusion be extended to consider retention at higher levels of significance. In practice however, there is only one sub-sample for which the 5% level cut-off is varied. For non-economic immigrants in LSIA1, EngBack (immigrants from the U.K., U.S.A., Canada, or Ireland) is maintained although the *t*-statistic is 1.466 (15% level of significance) as the inclusion results in an otherwise well-fitting model. For all other models, exclusion of variables in the general-to-specific process is unambiguous—*t*-statistics are either well below about 1.0 or well above 2.0 (in most cases variables are significant at the 1 per cent level or better).

LSIA1 Data Issues

Notwithstanding the preference for the general-to-specific process, some problems are encountered when applied to the LSIA1 data (but not the LSIA2 data): estimation problems are encountered (generally, failure to converge) when the model specification includes the full set of causal indicators for three waves. Thus, the general-to-specific approach forms the basis of model specification and model reduction in LSIA1, but in a number of cases the process must deviate from a strict application. For example, for NLF immigrants, inclusion of all causal variables (i.e. the general specification) causes failure to converge. Investigation shows that the variables causing the problems are either wave 1 time-invariant causal variables—or time-variant variables with very high correlations across time (see the discussion below regarding the treatment of such variables). To overcome this problem, an iterative process is used to establish which causal variables are causing model failure. When the offending variable(s) are identified, a two-step variant of the general-to-specific method is used: first, the general specification is re-estimated with necessary exclusions of wave 1 variables (with all wave 2 and 3 variables included). Second, an alternative general specification is estimated in which the previously excluded (offending) variables are included with as many other causal wave 1 variables as allowable for a model solution. In this way, the relative statistical significance (i.e. the *t*-statistic) for each variable can be considered and the general-to-specific method can be re-introduced.[6]

It is also important to model building to consider the across-time correlations for time-variant causal variables. Specifically, due to very high correlation between wave 1 and 2 (and 3) for some variables, a number of potentially time-variant indicators can only be included in one wave (usually, but not always, wave 1): that is, they are treated as if time-invariant. For example, the

[6] In some cases, a specific variable cannot be included in initial general specifications. When this happens, the variables are introduced into the process as soon as a solution can be obtained—but note that in no case did a re-introduced variable remain in the model through to the conclusion of the general-to-specific process.

correlations between Marstat (marital status) at waves 1, 2 and 3 in LSIA1 are 0.98, 0.90 and 0.95, and the correlations between Educat (education level) is 0.99, 0.99, and 0.99 (with similar values between waves 1 and 2 in LSIA2). "If very high correlations (e.g., $r > 0.85$) do not cause an SEM computer program to "crash" or yield a non-admissible solution, then extreme multicollinearity may cause the results to be statistically unstable" (Kline 2005, p.319). Thus, slow-changing time-variant variables (with resulting high across-time correlations) are included in only one wave (generally, but not always, wave 1—see below). This treatment is consistent with the underlying ideas relating to the linked MIMIC model discussed above: time-invariant and slow-change time-variant variables are viewed as "history", the value of SucSet2, depends on concurrent factors (time 2 causal variables), and on the latent state at time 1 (SucSet1) which is influenced by time 1 causal variables. Thus, the impact of time-invariant and slow-changing time-variant variables on SucSet2 mediates through SucSet1 (De Leeuw *et al.* 1997; Montfort and Bijleveld 2004).

When data problems restrict inclusion of some variables, the preferable option is to include time-invariant or slow-changing variables at time 1; in practice, there are cases in LSIA1 when this causes convergence problems, but there are no obvious reasons for the failure. Such cases are resolved pragmatically by allowing the time-invariant variable to enter at wave 1 and the slow-change variable at wave 2 (or wave 3).

Model Issues and Solutions LSIA1 and LSIA2

As noted previously, several time-invariant variables are very slow changing and hence their correlation between waves is very high, generally precluding their use in more than one wave. For example, the across wave correlations for the English language ability index (ELAI) are very high (i.e. $r > 0.95$), and so the ELAI can only be included in one wave in the general specification (but, as discussed below, ELAI is significant in only one model).

The inclusion of EngBack (immigrants from the U.K., U.S.A., Canada, or Ireland) and the PDI (Power Distance Indicator) causes model failure due to high correlation ($r > 0.80$) in LSIA1 for all sub-samples (but lower correlations in LSIA2 allow the inclusion of both). Following the procedure discussed above, models with EngBack and PDI are compared at early (general) stages of model building to establish which of the two is more informative for LSIA1.

Similarly in LSIA1, for non-economic and NLF immigrants, the inclusion of both Person (i.e. PA or MU) and Gender caused estimation problems (the tetrachoric correlations are above

0.70).[7] Likewise, for economic immigrants in LSIA1 the correlation between Person and Marstat (marital status) is 0.71, which causes model failure if both are included in the general model specifications. In these, and similar cases, alternative general model specifications were examined to suggest which of two conflicting variables should be included in the general specification.[8]

For non-economic immigrants in LSIA1 and LSIA2, and NLF immigrants in LSIA1, the inclusion of Wealth at wave 1 (Wealth1) in the general model causes non-convergence. When examined in a model with no other time-variant variables it appears to be unimportant (e.g. for non-economic immigrants in LSIA1 it is non-significant (coefficient = 0.066, *t*-statistic = 0.164)). In some cases a model can be estimated with Wealth2 (but excluding Wealth1), but model statistics point to problems (notwithstanding that when included in initial models it is statistically significant).[9] This result is not unexpected: wealth data are probably unreliable (suffering the same reporting problems as income data). In addition, in C2W1, 82 per cent (and 76% in C2W2) of immigrants report no wealth, and distributions are skewed by several very high values. Excluding wealth from models in which its inclusion is problematic suggests little is lost for this analysis (in models in which it can be included, it is not retained through the general-to-specific process). Nonetheless, future data collections may consider improving these data as a case has been made that wealth influences subjective well-being.

Other instances of sub-sample problems are NumAdult2 (the number of adults in the household at wave 2) for economic immigrants in LSIA1 (the reason for this is unclear, correlation between waves for NumAdult do not exceed 0.52), and relative income (Relinc) at wave 2 and 3 (possibly due to correlations between waves which range between 0.46 and 0.68). In some cases CameEco and CameFam cannot be simultaneously included (the reason is not clear, the two measures do not appear to be highly correlated—but in models where both can be included, in no case do both remain in the specific model). Age causes estimation problems when included in the general model for non-economic immigrants in LSIA2; examination of the measure on its own provides no guidance to the cause of failure, but in early steps in the general-to-specific process in which it was inserted and a model solution was obtained its coefficient was small and it was non-

[7] The use of interaction dummy variables (PA-males, PA-females, MU-males, and MU-females) did not solve this problem.

[8] For example, for NLF immigrants, the initial general specification included Gender, after some steps in the general-to-specific process Person was included in the specification successfully. For non-economic immigrants Gender entered the initial general specification, but was excluded through the general-to-specific process, but Person entered successfully.

[9] For example, the CFI test statistic is 1.0 for a perfect fit—but other statistics contradict this.

significant (e.g. the coefficient of -0.199 with *t*-statistic of -0.714).[10] Generally, age appears to be unimportant for the settlement process for LSIA immigrants, except for NLF immigrants (see below).

One final model issue requires attention. In all models except one, modelling, except for data issues outlined above, is reasonably straightforward producing sensible models (e.g. model solutions are considered proper as there are no out of range parameters such as negative variance estimates). For LSIA2 non-economic immigrants, however, models consistently estimate a negative error variance for RightMig. Following the literature (e.g. Byrne 1998; Brown 2006), this improper solution is overcome by fixing the error variance of RightMig to a very small positive value. The impact of this adjustment is small changes in estimated values, but the changes do not influence conclusions drawn. Brown (2006) notes that this is the most common form of improper solution (suggesting that one solution may be to "…collect a larger sample…" (Brown 2006, p.189) which is not possible in this case).

Thus data problems, encountered generally due to high correlations and common when analysing predominantly ordinal data, are dealt with on a case be case basis—guided by the general-to-specific model-discovery process.

MIMIC Model Results

Assessing MIMIC Model Goodness-of-Fit

Table 9 (LSIA1) and Table 10 (LSIA2—over) provides goodness-of-fit statistics for MIMIC models (for the three groups of immigrants) comparing the general model to the preferred model derived from application of the general-to-specific process. For both LSIA1 and LSIA2, the model statistics for the derived "specific" model indicate that the models are, at least, reasonable fits to the data, with most being a "good" fit.

[10] Note that in the later discussion the non-economic immigrant sub-sample for LSIA1 resulted in the least successful model.

Table 9: MIMIC Model of Successful Settlement—Goodness-of-Fit Statistics (LSIA1)

	Economic		Non-Economic		NLF	
	General	**Specific**	**General**	**Specific**	**General**	**Specific**
RMSEA	0.055	0.000	0.037	0.045	0.070	0.062
SRMR	0.049	0.058	0.059	0.063	0.059	0.079
GFI	0.996	0.992	0.999	0.971	0.998	0.979
AGFI	0.989	0.986	0.998	0.954	0.995	0.962
CFI	1.000	1.000	0.978	0.974	1.000	1.000
PGFI	0.407	0.577	0.514	0.599	0.431	0.535
PNFI	0.429	0.630	0.533	0.654	0.454	0.587
AIC	2762.8	322.0	1717.9	1207.0	3552.9	1719.6
AIC Null	27598.9	14639.7	41685.2	31156.7	32823.6	19781.6
CAIC	6317.5	1272.7	3538.5	2023.7	6663.2	3022.7
CAIC Null	27852.8	14787.3	41898.0	31307.7	33069.5	19953.7
ECVI	2.774	0.488	0.661	0.465	2.800	1.355
ECVI Null	27.710	14.699	16.051	11.997	25.866	15.588
Sample size	997	997	2598	2598	1271	1271

Notes: (1) General represents the initial general model results; Specific represents the model from the general-to-specific model reduction process (see text). (2) Data are weighted. (3) Estimation method is DWLS.

For LSIA1, all groups (economic, non-economic, and NLF) models have an RMSEA < 0.08 thus at least a "good" fit to the data; the SRMR < 0.10 is at least a "favourable" fit; the GFI, AGFI and CFI are all well above the 0.90 "good" fit requirement—statistics are generally little different to the value for the general (over-fitted) model; the parsimony index (PGFI) is greater in the specific model—indicating a preferred specification, as are the AIC, CAIC, and ECVI—which are also well below the null model values.

Table 10: MIMIC Model of Successful Settlement—Goodness-of-Fit Statistics (LSIA2)

	Economic		Non-Economic		NLF	
	General	Specific	General	Specific	General	Specific
RMSEA	0.056	0.052	0.098	0.099	0.102	0.070
SRMR	0.186	0.058	0.091	0.101	0.087	0.118
GFI	0.998	0.994	0.996	0.987	0.999	0.990
AGFI	0.995	0.987	0.989	0.976	0.980	0.981
CFI	1.000	0.967	1.000	0.962	1.000	0.952
PGFI	0.336	0.483	0.338	0.520	0.328	0.539
PNFI	0.368	0.515	0.360	0.560	0.348	0.582
AIC	1891.0	707.1	3034.9	1204.5	3328.7	582.2
AIC Null	30805.9	12120.0	28105.2	8851.1	26968.8	8157.0
CAIC	4361.9	1382.7	5379.9	1688.2	5852.3	952.1
CAIC Null	31018.6	12245.1	28308.0	8958.5	27177.6	8252.5
ECVI	1.336	0.500	2.863	1.137	3.143	0.550
ECVI Null	0.841	0.297	1.123	0.323	1.190	0.257
Sample Size	1416	1416	1062	1062	1060	1060

Notes (1) General represents the initial general model results; Specific represents the model from the general-to-specific model reduction process. (2) Data are weighted. (3) Estimation method is DWLS.

LSIA2 results are similar to LSIA1—except the specific model for non-economic immigrants has a RMSEA of 0.099 < 0.10 suggesting a “mediocre” model (noting that the RMSEA for the general model is about the same, 0.098). Models for economic and NLF in LSIA2 are “good” according to the RMSEA. Other model assessment statistics indicate at least a good fit to the data (i.e. the GFI, AGFI, and CFI are all well above the 0.90 “good” fit value), and model comparison statistics (AIC, CAIC, and ECVI) are all lower than the general model and null model.

In summary, group models (economic, non-economic and NLF) for LSIA1 and LSIA2 have at least satisfactory goodness-of-fit statistics, and in all cases, the parsimonious specific model is preferred to the over-fitted general model according to model comparison statistics. The MIMIC models can be considered successful in fitting models to data—the discrepancy between the theoretical and observed relations is not too large (Boomsma 2000). Thus, the influences on the evolution of immigrants’ successful settlement can be assessed.

Table 11 below provides details of the preferred specific MIMIC models of successful settlement for the three groups of immigrants, for LSIA1 and LSIA2, showing:

- Structural coefficients—the influence of settlement history on current outcomes for waves 2 and 3.

- Measurement model path coefficients (or factor loadings, λ,) (the **Λ** matrix)—the representation of successful settlement through reflective indicators.
- Formative model path coefficients (λ) (the Γ vector)—the influences on successful settlement of immigrant attributes.
- Variance of the latent variable SucSet at two or three points in time.
- The following formative indicators are not statistically significant in either cohort for any group of immigrants (and so are excluded from the table): Ownhome1, Ownhome3, Wealth1, Wealth2, NumAdult1 AttEng, CameEco, ELAI1, ELAI3, and Pension.[11]

[11] Pension is a dummy variable set to 1 if the immigrant is in receipt of any type of government pension (and zero otherwise). This indicator plays the same role as Umpbfit in the econometric analysis of labour market success—i.e. it represents the concept that government support provides financial assistance, but may also contribute to feelings of security and acceptance and hence is expected to contribute to successful settlement. It was not found to be statistically significant in any model for any group of immigrants (based on questions in Section U of the LSIA).

Table 11: MIMIC Model of Successful Settlement (LSIA)

	LSIA1 Economic	Non-Economic	NLF	LSIA2 Economic	Non-Economic	NLF
Structural Model (β) and Statistical Significance						
SucSet1→2	**0.660******	**0.617******	**0.713******	**0.610******	**0.457******	**0.562******
SucSet2→3	0.694****	0.636****	0.712****	n.i.	n.i.	n.i.
99% CI β_2 & β_3 Coincide	yes	yes	yes	n.i.	n.i.	n.i.
Measurement Model (Path Coefficients or Factor Loadings, λ) and Statistical Significance						
Encore	0.668****	0.499****	0.639****	0.913****	0.578****	0.727****
GHQ	0.604****	0.616****	0.785****	0.741****	0.721****	0.798****
LifeOk	1.000 (n.a.)	1.000 (n.a.)	1.000 (n.a.)	1.000 (n.a.)	1.000 (n.a.)	1.000 (n.a.)
RightMig	1.074****	0.989****	1.379****	1.322****	1.440****	1.184****
MIMIC Formative Model (Path Coefficients, γ) and Statistical Significance						
Person			-0.356****			
Gender			0.415****			
LMSI1	0.212****		n.i.	0.237****	0.387****	n.i.
LMSI2	0.107****		n.i.	0.073**	0.061****	n.i.
LMSI3		0.057****	n.i.	n.i.	n.i.	n.i.
Health1	0.303****	0.297****	0.405****	0.171****	0.361***	0.338****
Health2	0.075***	0.138****	0.073***			0.227****
Health3	0.109****	0.202****	0.136****	n.i.	n.i.	n.i.
OwnHome2			-0.104****			
EngBack1		0.054#		0.101****	1.225****	0.354****
PDI			-0.235****		-0.887**	-0.135***
MarStat1				0.086**	0.237****	
BetHome2	0.085***					
BetHome3	0.114***			n.i.	n.i.	n.i.
BetOff2	0.175****		0.091****	0.121****		

	LSIA1 Economic	Non-Economic	NLF	LSIA2 Economic	Non-Economic	NLF
BetOff3	0.198****		0.108****	n.i.	n.i.	n.i.
Age			0.094**			0.118****
Timeoz1		-0.061***		-0.084****	-0.107***	
TimeOz2		0.064****	0.089****	0.065****	0.050***	
TimeOz3		-0.077****	-0.130****	n.i.	n.i.	n.i.
NumAdult2			0.197****	0.060****		
NumAdult3	-0.050***		-0.104****	n.i.	n.i.	n.i.
Educat1	-0.251****	-0.291****	-0.232****	-0.169****		
Relinc1						0.116****
Relinc2						0.027****
Relinc3			0.064****	n.i.	n.i.	n.i.
Spons		-0.140****			-1.123****	
CameFam	0.084**			0.137**	0.766****	
ELAI2	0.044***					
Variance of Successful Settlement						
Variance SucSet1	0.687	0.676	0.432	0.519	0.539	0.463
Variance SucSet2	0.625	0.604	0.372	0.467	0.483	0.349
Variance SucSet3	0.651	0.678	0.408	n.i.	n.i.	n.i.

Notes: (1) Estimation method is DWLS. (2) Data are weighted. (2) n.a. (not applicable) indicates a *t*-statistic (based on the standard error) is not available for the “fixed” reference variable. (3) n.i. (not included), i.e. wave 3 variables in LSIA1, and the LMSI for NLF immigrants (index is zero for all NLF immigrants). (4) Statistical significance levels: **** = 0.1%, *** = 1%, ** = 5%, * = 10%, # 15%. (5) Equality of β coefficients across time based on calculated 99% confidence interval—as Cov(β_2 ,β_3) not estimated. (6) Excluded formative indicators (non-statistically significant for all groups in both cohorts are: Ownhome1, Ownhome3, Wealth1, Wealth2, NumAdult1 AttEng, CameEco, ELAI1, ELAI3, Pension.

Reliability of the MIMIC Model

Before discussing the implications of the MIMIC models for successful settlement it is useful to consider the overall reliability of the model

An informal, method of judging the reliability of the MIMIC model is to compare the MIMIC model with the less complex longitudinal SEM of SucSet. This comparison suggests the MIMIC model is reliable based on the following observations:

- The measurement model in the MIMIC specification (i.e. the model of reflective indicators) is congruent with the longitudinal SEM for SucSet for immigrants—statistically significant factor loadings on Encore, GHQ, LifeOk, and RightMig retain their relativity (e.g. the loading for RightMig is highest, and for GHQ is lowest, in the longitudinal SEM and MIMIC models).
- The structural coefficient (β, between SucSet at different waves) is comparable across models (e.g. in the longitudinal SEM for LSIA1 it ranges between 0.663 (non-economic, wave 1 to 2) and 0.837 (NLF, wave 1 to 2) and in the MIMIC model it ranges between 0.617 (non-economic, wave 1 to 2) and 0.713 (NLF wave 1 to 2)).
- The estimated variance of SucSet decreases between waves 1 and 2 (all groups), but increases between waves 2 and 3 in LSIA1 in both approaches (e.g. for economic immigrants in the longitudinal SEM the variance of SucSet is 0.685, 0.624 and 0.656 compared to 0.650, 0.609, and 0.651 in the MIMIC model).
- The measure of indicator reliability (the SMC) for individual reflective indicators are similar in both approaches (e.g. for economic immigrants in LSIA2 the longitudinal SEM value for SucSet is 0.49 compared to 0.55 in the MIMIC model).

Thus, while the expectation is that the two methods should give different estimates of coefficients (since the MIMIC model accounts for both formative and reflective effects), there are no unexplained extreme divergences—which, if present, would be a cause for concern.

Interpreting the MIMIC Model

Statistics indicate a successful MIMIC model has been obtained. Thus, the evolution of successful settlement can be assessed: the unobserved latent variable SucSet can be modelled with “cause” and “effect” indicators (noting the limitation of all models, causality is not proven, but relies on the underling model rationale—see below).

Variance Explained—Model Reliability

The reliability of the unobserved latent construct, SucSet, can be assessed by considering the implied proportion of variance in SucSet that is explained by the model—that is, the reliability measures how well the reflective indicators serve as instruments for SucSet (Jöreskog and Sörbom 2001).[12] Before considering this approach note however that reliability does not provide a measure of model goodness-of-fit (Hayduk 1996; Kelloway 1998) as it does not provide information about whether the model reproduces the observed correlation matrix;[13] once model goodness-of-fit is established (see above) reliability can be considered.

The reliability for the latent constructs range from the minimum of 0.20 for non-economic immigrants at C1W1 to a maximum of 0.75 for NLF immigrants at C1W2). That is for example, collectively the formative (causal) indicators are able to predict 75 per cent of the (model implied) variation in SucSet2 for NLF immigrant in LSIA1 (Brown 2006) (see Table 12). The measure of reliability shows that the proportion of variance explained varies, but lower values are not unexpected given that much of the data are (as noted above) dichotomous or ordinal, the exploratory nature of the models, and the opportunistic use of the data (which was not collected with the intention of applying sophisticated statistical techniques. On the other hand, higher reliability values suggest a remarkable degree of accuracy given these conditions.

Table 12: Reliability—MIMIC Models (LSIA)

	LSIA1			LSIA2		
	Economic	**Non-economic**	**NLF**	**Economic**	**Non-economic**	**NLF**
SucSet1	0.277	0.203	0.466	0.250	0.221	0.535
SucSet2	0.638	0.493	0.745	0.507	0.361	0.671
SucSet3	0.607	0.488	0.641	n.i.	n.i.	n.i.

Notes: (1) Reliability is the proportion of the (model estimated) variance in the latent variable explained by the reflective indicators. (2) n.i. represents not included in the model (i.e. no wave 3 for LSIA2).

The Measurement Model

Interpreting the MIMIC measurement model is identical to interpretation of CFA or longitudinal SEMs. Path coefficients (factor loadings, λ,) are the effect of the latent variable on its reflective indicator (i.e. the extent to which SucSet is jointly reflected in Encore, GHQ, LifeOk, and RightMig).

[12] SucSet is unobserved, but an estimate of its variance is a model output (which does not require an estimate of the latent variable itself).

[13] That is, the calculation of the proportion of variance of the unobserved latent variable (SucSet) does not refer to the original data, i.e. "There is nothing in the formula for [reliability] that refers to the data, or the match of the model's implications to the data" (Hayduk 1996, p.215).

The Structural Model

In the structural model, statistically significant coefficients (β) between SucSet across waves represents the dependence of SucSet at wave 2 on, amongst other things, the value at wave 1 (e.g. SucSet2 can be predicted to some extent by SucSet1—and similarly for SucSet3 in LSIA1). Thus, the coefficient has the usual regression model interpretation. In MIMIC models, the βs are strongly significant, and relatively large (ranging from a minimum of 0.457 for non-economic immigrants in LSIA2 to a maximum of 0.713 for NLF immigrants in LSIA1)—as with subjective well-being, SucSet exhibits homeostasis. The structural coefficients appear stable, the 99% confidence interval for β_{12} and β_{23}, in LSIA1 for the three groups of immigrants, coincide implying stability of the relationship between SucSet at various points in time.

The Formative (Causal) Path Coefficients

Coefficients for the formative indicators are interpreted as are coefficients in regression analysis.[14] Thus, for example, all other things equal, a statistically significant positive coefficient on a dichotomous indicator implies a higher mean value of SucSet for the group of immigrants identified by the formative indicator (Brown 2006).

MIMIC Model—Influences on Successful Settlement

The empirical results for the MIMIC models show that the formative indicators that theoretically, or intuitively, are thought to influence successful settlement, are statistically significant for at least one group of immigrants (i.e. economic, non-economic, or NLF) in one or more waves of the LSIA.[15]

Importantly, labour market outcome (measured as the LMSI) was statistically significant, in at least one wave, for all immigrants in the labour force. In LSIA2, the LMSI was significant in both waves for economic and non-economic immigrants: in C2W1, it was more important for non-economic immigrants, in C2W2, marginally more important for economic immigrants. The impact at wave 2 was significantly reduced—for economic immigrants, the coefficient fell from 0.237 to 0.073, and from 0.387 to 0.061 for non-economic immigrants. A similar pattern is observed for economic immigrants in LSIA1 between waves 1 and 2, but the LMSI

[14] Noting that in comparison with an econometric model approach, the MIMIC model allows testing complex models whereas multiple regression "…would provide only separate "mini-tests" of model components that are conducted on an equation-by-equation basis" Tomarken and Waller (2005, p.34).

[15] Three indicators were found not to be statistically significant (CameEco, Wealth, and Pension), but their roles are an additional—or complementary—aspect of a dimension covered by another indicator (e.g. CameEco and CameFam).

was not significant at wave 3—and for non-economic immigrants the LMSI was significant only at wave 3 (with a coefficient of 0.057—small and comparable with wave 2 for LSIA2). Thus, it is clear that, generally, labour market success is much more important to immigrants shortly after arrival than later.[16] Moreover, coefficients for the LMSI,[17] even at wave 1, are not particularly large—the labour market matters, but, for example in LSIA1, less so than health. Thus, while labour market success is expected to influences SucSet, it does not define SucSet (even for economic immigrants). Moreover, the 95 per cent confidence interval for the relatively small LMSI coefficients for economic and non-economic immigrants in LSIA2 overlap and thus the labour market can be interpreted as being equally important for economic and non-economic immigrants.[18]

English language ability (measured as the ELAI) influenced outcomes only for economic immigrants in LSIA1. There are three explanations for this outcome. First, the relationship between ELAI and country of origin suggests that EngBack i.e. if from U.S.A., U.K., Canada, N.Z., or Ireland) supplants English language ability as an explanatory influence (e.g. EngBack was significant for all immigrants in LSIA2, with a coefficient ranging from 0.101 for economic immigrants to 1.225 for non-economic immigrants). Thus, either English language ability or EngBack mattered, but not both—either measure captures the impact of "Englishness". Second, expectations may play a role—those with well-developed English language ability have expectations fulfilled as they are more likely to settle quickly, those with less well-developed abilities settle less quickly, but this simply fulfils expectations. Third, English language ability influences labour market success (i.e. through the LMSI) and hence the impact on SucSet may be indirect (as with education discussed below).

An interesting result emerges from the measure of cultural similarity (the PDI). Excluding those from U.S.A., U.K., Canada, N.Z., or Ireland, cultural similarity with Australia did not improve SucSet. The reason for this may be, as with labour market outcome, and English ability above, due to expectations. Thus, immigrants from a closer cultural background (excluding "Englishness" itself) may expect to settle readily in Australia, but find that being an immigrant can be difficult. On the other hand, an immigrant from say the Sudan who

[16] The correlation between LMSI1 and LMSI2 in LSIA2 is 0.43 for non-economic immigrants and 0.40 for economic immigrants, and correlations for LSIA1 are 0.46 (C1W1 to C1W2) and 0.37 (C1W2 to C1W3) for economic immigrants, and 0.37 and 0.40 for non-economic immigrants. Thus, early values of the LMSI are probably not a "history" variable at later waves.

[17] As noted previously, the LMSI is rescaled to a range of zero to one.

[18] Note that in C2W2, only 4% of economic immigrants and 16% of non-economic immigrants were unemployed, which explains why a separate dummy variable for labour force status included experimentally in the models was not significant.

knows Australia is foreign, finds Australia to be foreign but is not surprised and has come to Australia prepared to be different. This is an area where qualitative information would be useful—questions relating to expectations would be a useful addition to future surveys of immigrants.

The impact of Education on SucSet was not as expected. More education (measured on a scale of 1 (representing six or less years of school) to 9 (higher degree) is associated with lower SucSet for all immigrants in LSIA1 and for economic immigrants in LSIA2 (but not significant for other LSIA2 immigrants). There are two possible explanations for this result. First, the result is skewed because immigrants in the middle of the education ranking are those with a trade or technical/professional qualification who found employment and acceptance (including social participation) into mainstream Australia facilitated by trade or professional connections. Thus, for some immigrants, education (and/or use of qualifications) influences labour market success (i.e. the LMSI) and hence the impact of education is indirect. Similarly, many post graduate or higher-degree qualified immigrants struggle to find employment at their professional level (Birrell *et al.* 2007) and perhaps this leads to less positive attitudes to successful settlement (and hampers social participation). Second, those with lower educational qualifications held more easily fulfilled expectations about their settlement process, highly educated individuals may have assumed that education equated with ability to settle, but this may not have been the case. As with the impact of the PDI, this is an area where qualitative information would be useful. This paper cannot model adjustments to expectations that may influence the settlement process, nor is there a way to model social withdrawal or feelings of rejection (due to overt or covert discrimination.

The impact of time in Australia (TimeOz) is also interesting. Where statistically significant (all immigrants except economic in LSIA1, and NLF in LSIA2), TimeOz1 (average 5 months in Australia) has a negative coefficient, TimeOz2 (average 17 months) a positive coefficient, and in LSIA1 TimeOz3 (average 41 months) reverts to a negative coefficient. TimeOz could be interpreted as showing that immigrants, on average, have some difficulty in becoming acquainted with Australia's norms and customs; after a year or so they felt things are going well (all things considered), but after 3 years the honeymoon is over (whether this is the case for LSIA2 is beyond these data).

Sponsorship was only statistically significant for non-economic immigrants (e.g. in LSIA2, about 46% of sponsored immigrants are non-economic and 41% are NLF), and in both cohorts it was associated with a lower level of SucSet. The expectation was that sponsorship would improve social participation, but it appears that when relevant it delays wider participation (perhaps by providing a closer association with the sponsor, but not with the Australian community in general (e.g. in LSIA2 about 75% of sponsors were a husband, wife, or fiancé).

On the other hand, immigrants who indicated that they came to Australia specifically to join family in Australia or to get married (CameFam) had higher levels of SucSet (CameFam was unevenly distributed across all three groups of immigrants, e.g. in LSIA2, 15% of economic, 66% of non-economic and 46% of NLF immigrants). In this case social participation appears to have been enhanced as expected. CameEco (i.e. immigrants who came for explicit economic reasons as opposed to reasons such as better life-style) do not appear to have settled any more successfully than others.

Health matters to all immigrants, as expected its impact is positive. In LSIA1, health is an explanatory variable at all waves, for LSIA2 only at wave 1 for economic and non-economic immigrants, perhaps entering as "history" at wave 2;[19] or for immigrants subject to a more stringent entry requirements (LSIA2), poor health was less common.[20]

A number of variables influence SucSet for specific groups at specific times and are congruent with expectations:

- In C1W1 for NLF immigrants Person (PA versus MU) and Gender are significant: being a PA tends to reduce SucSet—consistent with responsibilities of the household head; and being female reduces SucSet—consistent with the idea that the female (usually also the MU) generally stayed at home and hence was somewhat isolated. Interestingly Person and Gender do not influence other immigrants in any wave: but for Gender, this is consistent with findings that suggest that it does not directly influence subjective well-being.
- Marital status (Marstat) is significant in LSIA2 (there is a positive coefficient for all but NLF immigrants[21])—having a partner provides support and facilitates settlement

[19] The correlations for health between C2W1 and C2W2 is 0.58 (0.56) for economic (NLF) immigrants, which suggests it may be a "history" variable at wave 2, but it is 0.37 for non-economic immigrants so this does not explain its absence in C2W2 for this group (in LSIA1 correlations range from 0.34 to 0.49). Using the alternative measure of physical health, DrVisit, did not alter the result to any important degree.

[20] For example, general skilled immigrants must provide evidence of a recent health examination for all family members included in the visa application.

[21] The impact of a partner on NLF immigrants is assumed positive. As most NLF immigrants have a partner, there is no impact of the dummy variable for marital status (Marstat) in the MIMIC analysis. For example, in LSIA1 81% of the NLF immigrants (all MUs and 72% of PAs); and in LSIA2, 87% of NLF immigrants (97% for MUs and 81% for PAs), had a partner. Since NLF immigrants have less outside contact than immigrants in

(and for those in the labour force, provides an incentive to succeed in the labour market thus indirectly influencing SucSet).

- Having a better home (BetHome) or being financially better off (BetOff) at wave 2 or 3 (variables representing financial success) are more important in LSIA1 than LSIA2. BetOff is significant for economic immigrants in LSIA2, and for NLF immigrants in LSIA1, but BetHome influenced economic immigrants in LSIA1. On the other hand, the proxy for relative income mattered only to NLF immigrants in LSIA2 (perhaps changes in relative income were not sufficient in the short to medium period covered by the LISA to be important). On average, measures of financial success matter more in LSIA1—perhaps because for LSIA2 entry requirement resulted in financial outcomes being more homogenous.
- Higher age is associated with greater SucSet for NLF immigrants (but not relevant otherwise). This is consistent with the U-shaped relationship between subjective well-being and age): NLF immigrants are on average about 5 years older than other immigrants (with a standard deviation of about 16 years compared to about 10 years or less for others).

In two cases, variables influence SucSet for a group in one wave and not as expected; speculation can suggest causes, but further investigation in follow-up studies may be warranted:

- Home ownership (OwnHome) reduces SucSet at C1W2 for NLF immigrants, but applies to only 4 per cent of the sample (otherwise the variable is not relevant).
- The number of adults (NumAdult) in the household increases SucSet for NLF immigrants in C1W2 and economic immigrants in C2W2, but reduces SucSet for C1W3 economic and NLF immigrants. Thus, at wave 2 the variable acts as expected and facilitates social participation and support. In wave 3 however, the negative impact suggests perhaps pressure of additional adult household members (cost of living or lack of privacy) reduced SucSet.

---

the labour force it is assumed that the contact with their partner does tend to improve the settlement process. An alternative explanation is that there are asymmetric impacts of having a partner. For example, labour force participants find settlement enhanced by a partner at home who provides support outside work (or seeking work), but the at home partner does not feel support from the non-NLF partner. The LSIA data for NLF immigrants are not sufficiently diverse (between partnered or not) to indicate which explanation is more likely but the positive influence for NLF of partners seems intuitive.

Changes in Successful Settlement

As noted previously, MIMIC models allow assessment of the evolution of successful settlement: the attraction of the MIMIC model is that the "causes" (or influences) on successful settlement can be assessed. Although a "factor score" for the latent construct, SucSet, can be generated, estimation of scores can be problematic. There are several ways to estimate a factor score, but no particular method is preferred, the various methods do not necessarily provide consistent results or ranking of individuals, and hence care is required when comparing factor scores for individuals (Brown 2006). On the other hand, estimates of average group values result in less ambiguity, as does comparing averages across time. Noting this caveat, a "coarse" group-average score for SucSet can be generated by solving the combined structural and measurement components of the MIMIC model.[22] That is, the estimated measurement path coefficients (or factor loadings, λ) are applied to the rescaled[23] group-average values of the reflective indicators; the history of SucSet is included by applying the structural coefficient (β) to the group-average value of SucSet at previous waves.[24]). Table 13 provides estimates of SucSet for each wave of the LSIA (rescaled so that the maximum is 100 at wave 1[25]) for the MIMIC model.[26]

---

[22] Equivalently, for the MIMIC model, an approximate value of SucSet can be generated by evaluation of the model at the mean value of causal (rescaled) variables weighted by estimated coefficients in the formative (causal) model.

[23] The *NX* transformation is used to rescale the reflective indicators before applying the factor loading to calculate a "factor score". The transformation $NX = (x_i - \min(x)) / (\max(x) - \min(x))$ is used to deal with different scales in the reflective indicators (*NX* rescaled indicators have a range of [0:1]. Rescaling is applied to the pooled data (i.e. waves are stacked). In addition, to ensure rescaling does not alter the relationship between groups (i.e. economic, non-economic, and NLF) rescaling is applied to combined data. A "factor score" is then calculated on wave and group specific sub-samples. Rescaling also contributes to the resulting score being "coarse" as ordinal data are treated as if continuous (given that reflective indicators are ordinal measures representing the unobserved underlying continuous variable, evaluation at the mean of the (rescaled) observed indicator is a coarse representation of the underlying continuous measure). Note, standardised data (mean zero unit standard deviation) cannot be used as this leaves the reflective indicators with different scales.

[24]E.g. $SucSet_2 = \beta * SucSet_1 + \lambda_1 * Encore_2 + \lambda_2 RightMig_2 + \lambda_3 LifeOk + \lambda_4 GHQ$ is a coarse proxy of SucSet since measurement errors in the latent variables and the reflective indicators are ignored (although since the errors are assumed mean zero, evaluation of this expression at the mean of measured variables reduces the coarseness somewhat).

[25] That is, for four *NX* rescaled indicators the maximum total is 4, for convenience this has been rescaled to 100 (divided by 4, multiplied by 100).

[26] Given the same method is used to generate a value for SucSet, values of SucSet can be compared across groups.

Table 13: MIMIC Model Based Factor Score for SucSet (LSIA)

	C1W1	C1W2	Change	C1W3	Change	C2W1	C2W2	Change
Economic	58	97	66%	96	-1%	58	94	62%
Non-Eco	54	86	60%	84	-2%	54	80	48%
NLF	54	91	69%	85	-6%	50	76	52%

Notes: (1) Non-Eco represents non-economic. (2) Score estimates are based on path coefficients (i.e. factor loadings, λ) and (rescaled using the *NX* transformation) reflective indicators—evaluated at the mean of the reflective indicators for groups at each wave, with the influence of the previous value of SucSet included (through the β structural coefficient).

Table 13 demonstrates that:

- SucSet increased between wave 1 and wave 2.
- Increases in SucSet between waves 1 and 2 were larger in LSIA1 than LSIA2—economic immigrants had the smallest difference.
- Non-economic immigrants recorded the smallest increases, and economic immigrants the largest, from wave 1 to 2.
- Immigrants in LSIA1 had higher values of SucSet at wave 2 than LSIA2 immigrants.
- SucSet did not improve between waves 2 and 3 for LSIA1 immigrants. Falls were recorded, but they may be too small to be significant (i.e. about equal for economic and non-economic immigrants at 1% to 2%, with a 6% fall for NLF immigrants—which may be large enough to signify a real fall).
- The falls in SucSet in LSIA1 immigrants do not suggest immigrants were not better settled at wave 3 than wave 1—but these data cannot indicate whether the trajectory continues downwards, or reaches a turning point.
- Economic immigrants start with a higher value for SucSet, and retain the advantage, but given that economic immigrants are selected for success in the labour market, their successful settlement advantage does not appear stark compared to immigrants who were not selected for labour market success.
- LSIA1 immigrants appear to be more successful at wave 2 than LSIA2 immigrants, and this is most obvious for NLF immigrants.
- In LSIA1 by wave 3 NLF and non-economic immigrants are equally as successful (in LSIA2 they are approximately equal at wave 2).

An important finding from this research is that the perception of immigrants in LSIA1 is that they were, at best, no better settled at wave 3 (after about 41 months in Australia) than wave 2 (after about 17 months). It is possible that by wave 3, immigrants felt that their immigration "honeymoon" was over. Whether the wave 3 result is a consistent result is unknown (i.e. there is no wave 3 for LSIA2), but the fall in SucSet is consistent with data remigration of immigrants after several years in Australia. Thus, the longer the time in Australia, the greater the outflow (e.g. in 2006-07, of the 35,221 permanent departures of overseas born persons, 14% had been in Australia for less than 2 years, 16% between 2 and 5 years, and 73% over 5 years (DIAC 2007b)).

Assessing the Settlement Process

Although a factor score can be estimated for SucSet, the main attraction of the MIMIC model is the ability to quantify the influence of causal indicators for particular groups of immigrants (in this case, specified by selection criteria and labour market activity). Thus, the MIMIC models suggest the attributes that increase, or decrease, the likelihood of successful settlement. This can therefore inform selection policy for successful settlement, and the after-arrival services that may improve the settlement outcome.

The discussion below is in terms of individual attributes, but it is noteworthy that at least one casual indicator from the settlement domains specified in Table 2 above plays a role in successful settlement. Thus, it is informative to describe successful settlement in terms of social participation (e.g. TimeOz), economic participation (e.g. the LMSI), economic well-being (e.g. BetOff) and health (Health).

MIMIC Model Implications for Immigrant Selection Criteria

The Labour Force

As noted above, economic immigrants do not appear appreciably more successfully settled than other immigrants—although if the change in SucSet followed the paths implied in Table 13 above, differences between economic, non-economic and NLF immigrants will increase as their paths diverge. Nonetheless, selection based on labour market readiness does not appear to lead to stark contrasts in successful settlement between economic and other immigrants in the short to medium term.

MIMIC model results above suggests that labour market success (the LMSI) is a more important factor for the settlement process for immigrants in the labour force in LSIA2 than LSIA1—this may be the result of changes in selection policy between LSIA1 and LSIA2 with a greater emphasis on job-readiness. The direct contribution to successful settlement of the LMSI is appreciably less at wave 2 than wave 1 (although wave 1 values influence wave 2 SucSet as history); it is less for economic immigrants than non-economic immigrants in C2W1, but somewhat larger for economic immigrants at C2W2; it is greater for economic immigrants in C1W1 and C1W2 (as it is not significant for non-economic immigrants).[27] For economic immigrants in LSIA1 and LSIA2, the labour market matters at waves 1 and 2, but in LSIA2 it also matters for non-economic immigrants. Thus, there is some evidence that the labour market is, generally, more important for economic than non-economic immigrants. In all cases, the direct impact of the labour market diminishes considerably by wave 2 (with a small impact for non-economic immigrants at wave 3). Labour market success matters to successful settlement—at some stage in the settlement process, to all immigrants in the labour force—but the labour market does not define successful settlement. For example in LSIA2, an English background has about four times the impact of the labour market at wave 1: and in LSIA1, health usually matters more than the labour market.

For immigrants in the labour market, labour market success contributes to successful settlement. Therefore, policies developed to improve the labour market outcome of labour market participants may have beneficial impacts on immigrant successful settlement. Indirectly, tightening of selection criteria for economic immigrants (i.e. for improved labour market success) could contribute to successful settlement, but global changes to demand for immigrants do not make this a practicable longer-term option.

Further, in relation to models for the LMSI, a number of avenues are available to improve labour market success of immigrants who are in Australia—policy need not be concerned solely with selecting job-ready immigrants from a global market for skilled immigrants that may restrict Australia's ability to select immigrants from a narrowly defined group. An alternative to selecting job-ready entrants is to select immigrants who could be job-ready after some investment by Australia such as training, or programs designed to help immigrants convert foreign qualifications to a standard acceptable to the Australian labour market.

Non-Labour Force Attributes

Immigrant attributes that are influenced by selection criteria that are relevant to all immigrants are health and "Englishness" (i.e. country of origin as English speaking country (EngBack) or English language ability (ELAI)).[28] This infers that no single immigrant selection criterion can be altered to improve successful settlement for all immigrants. Since immigrant selection policy, generally (and reasonably), excludes immigrants with health problems and immigrants

[27] For example, the average value of the LMSI in LSIA2 is 0.72 at C2W1 and 0.80 at C2W2 for economic immigrants; and 0.61 and 0.69 for C2W1 and C2W2 for non-economic immigrants. The coefficients on the LMSI in the LSIA2 are 0.237 for economic and 0.387 for non-economic immigrants at wave 1 compared to 0.073 for economic and 0.061 for non-economic immigrants at wave 2. Therefore, the simple (unscaled) impacts (i.e. coefficient by average group value for the LMSI at each wave) are therefore: for economic immigrants at wave 1, 0.17 versus 0.06 at wave 2; and for non-economic immigrants 0.24 at wave 1 versus 0.04 at wave 2.

[28] The exception is "Englishness" for NLF immigrants in LSIA1.

generally have good health,[29] there is little opportunity to influence successful settlement by tightening health requirements. Similarly, it is clear the settlement process is easier for English background immigrants (i.e. the U.K., Canada, the U.S.A. and Ireland), but it is generally these immigrants that are in demand in the other traditional immigrant countries of Canada, N.Z., and the U.S.A., and who are becoming important to non-traditional developed economies including the expanding European Union. Thus, restricting applicants further based on their "Englishness" is not a viable option—it may improve successful settlement for selected immigrants, but the potential supply will be restricted whereas increasing the potential supply of quality immigrants is required. An alternative approach is to assist immigrants' settlement by providing extended, accessible, English language tuition (and introductory social studies), in an environment that is conducive to forming social contacts, with assistance in attending (e.g. councils provide community buses for aged persons, community transport for immigrants could be viewed as investment in Australia's future).

Time in Australia (TimeOz) detracts from successful settlement at wave 1, enhances successful settlement at wave 2, and again detracts for C1W3. The settlement "honeymoon" appears to occur between wave 1 and 2, all other things equal, immigrants face more difficulties in their early settling in period, they then experience improvements as they become used to Australia, but (consistent with the fall in SucSet for C1W3, and the emigration referred to above) the "honeymoon" is over and the settlement process becomes more demanding. Time in Australia does not automatically enhance successful settlement. Thus, the MIMIC model results suggest that longer-term immigrant programs may have a place in assisting the settlement process, but further work is required to develop such programs—informed by qualitative research which may suggest features of Australia that are most difficult for immigrants to deal with.

No other single attribute is relevant to all (or most) immigrants' successful settlement. Nonetheless, the MIMIC model results are clear: immigrants with a partner are more successful than others (and this is more so for non-economic immigrants). This implies settlement can be enhanced by immigration policy that encourages immigrants with families—but not with the current restriction that points awarded towards the required total are only awarded for skilled partners. For example, selection policy could recognise the longer-term benefits to Australia of families with children by awarding extra points for a

[29] For example, the average (non-rescaled) value for health of NLF immigrants at C2W2 was 4.05 on a scale with maximum five.

partner and young children (immigrants make capital investments in Australia, Australia can reciprocate by investments in its future). Thus, although children (NumChild) make no direct contribution to SucSet (suggesting there is a balance between the pleasure and support they provide and their costs—including financial, time, and effort) healthy, young, immigrant children can become valued members of Australian society in later years—particularly if their parents are successful (Castle *et al*. 1998). Similarly, in the early period of settlement, additional adults in the household (NumAdult) can contribute to successful settlement. This suggests further consideration be given to the extended family that can be included in the immigration application. An adult accompanying immigrant, for whom the applicant is financially responsible, would make little demand on Australian resources, but could contribute to the settlement process of desirable family members. Moreover, as argued below, allowing extended family may contribute to the attractiveness of Australia as a destination in an increasingly competitive environment for quality immigrants. In addition, family reunion policies should also be seen as a strategy that will contribute to successful settlement in the longer-term. The role of family reunion programs as the residual of the total inflow after skilled and humanitarian programs are filled warrants reconsideration in light of their current and potential contribution to Australia's welfare; this could include added incentive to encourage those with young families. Immigrant selection policy must take a longer-term view than an excessive emphasis on job-ready immigrants—this may be one way Australia could increase its attractiveness in a world of increasing demand for quality immigrants.

Immigrant age is another selection criterion that requires reassessment. Interestingly, age appears to assist the settlement process for NLF immigrants (in both cohorts)—with no impact for others. This suggests that the current points system restriction on those aged 45 or under (when applying) should be reassessed—particularly in a world of increasing life span. Given current trends, an immigrant aged 50 could make a valuable contribution to Australia's workforce for 20 years or more. Moreover, older immigrants are more likely to have a partner, which will increase their likelihood of successful settlement—and their partner may enter the workforce and so contribute to Australia's economic welfare. In addition, if they have accompanying younger dependents, Australia's longer-term labour force prospects are improved. Moreover, a change in age policy by Australia may simply be the first in a country progression as others see the potential benefits.

Immigrants who come to Australia for family reasons (CameFam), other than NLF immigrants, appear to settle more successfully thus reinforcing the view that family migration and reunion programs should not be the residual of the skilled and humanitarian programs if successful settlement is of importance.

The influence of measures of financial success are interesting: in both cohorts at least one measure representing the financial domain (e.g. BetOff or Relinc) is significant for economic and NLF immigrants, but this domain does not appear to influence non-economic immigrants. In addition, several financial domain measures are significant in LSIA1, but few in LSIA2. Perhaps expectations drive this result for economic immigrants compared to non-economic immigrants. It is also likely that access to social security payments after 6 months for LSIA1 immigrants raised expectations regarding the financial domain compared to LSIA2 who could gain access only after 2 years. Nonetheless, coefficients were not particularly large, but become larger in later waves suggesting expectations adaptation.

LSIA1 versus LSIA2

As noted previously, significant changes to Australia's immigrant selection criteria for economic immigrants, and access to welfare payments, were introduced between LSIA1 and LSIA2. Thus, from the perspective of the settlement process, LSIA1 and LSIA2 provide a (short-term[30]) trial: what influence did changes to selection criteria for labour market success have on successful settlement of points tested immigrants?

The MIMIC models suggest that there are few differences in successful settlement evolution for comparable immigrants (i.e. waves 1 and 2). Group-average scores for successful settlement were not starkly different between cohorts (noting the caveat regarding score estimates above)—tighter selection criteria did not improve the rate of change in successful settlement. Thus, tighter selection criteria is not the way successful settlement can be positively influenced. As with other issues, perhaps immigrant expectations explain this result—a more arduous selection process raises expectations, which are harder to fulfil.

Evaluating the MIMIC Model Method

The information that comes out of models in this paper requires the application of advanced analytical techniques based on latent variables that may not be familiar to economists. Thus, an obvious question can be asked to evaluate the research in this paper:

> Have the empirical methods, which are complex and hence more demanding of the data and the analyst, added sufficiently to the understanding of successful settlement to warrant the introduction of the advanced techniques?

[30] That is, outcomes at wave 2 can be compared, but not wave 3.

To address this general question consider the following questions and answers relating to the advantages of the linked MIMIC models:

(1) Why not use a single measure to represent successful settlement in a common regression model? Measuring successful settlement by a single imprecise indicator variable (e.g. immigrant satisfaction with life in Australia), fails to incorporate measurement error, and is a coarse measure that fails to incorporate the multidimensional aspects of successful settlement. Thus, an immigrant may be conditionally (all other things equal) satisfied with their life in Australia, but all other things are not equal for all immigrants (e.g. some may be satisfied, but not be prepared to encourage others to immigrate to Australia). That is, no single indicator sufficiently captures the multiple aspects of successful settlement, and while it is not possible to capture all aspects of successful settlement, multiple indicators cover multiple attributes.

(2) Is cross-sectional factor analysis (FA) appropriate? Cross-sectional analysis does not track dynamics of the settlement experience. As with the benefits of longitudinal data econometric models versus cross-sectional analysis, longitudinal data can examine change, differentiate between change in cohorts and change for individuals and reduce the possibility that a particular cross-section is atypical. Moreover, for survey data, across-time correlations distort empirical estimates if ignored.

(3) Are longitudinal structural equation models (SEM) (i.e. factor analysis with a time dimension) sufficient? Longitudinal methods overcome the cross-section misspecification, but they only track the progress of successful settlement—they do not provide guidance or the reason why successful settlement follows a particular path. Thus, the addition of the formative component to the SEM to form the MIMIC model allows simultaneous assessment for cause (formative indicators) and effect (reflective indicators)—formative components allow determination of the different influences on the path of successful settlement.

(4) Why not use the reduced form of the MIMIC model (i.e. remove the latent variable for successful settlement and represent the underlying model in the more familiar econometric form)? It is well known that the MIMIC model can be respecified as the reduced form in which only observed variables (i.e. formative and reflective indicators) appear. This allows estimation (as a system of simultaneous equations) but the path coefficient for the formative (causal) indicators and reflective indicators cannot be

extracted—only a composite coefficient is available, unless it is assumed that all variables are measured without error—and thus nothing can be said about the relative importance of reflective and formative indicators.[31] As well as failing to incorporate measurement error, the reduced form system, (i) ignores across-time correlation between reflective indicator errors, and (ii) does not model correlation between formative indicators within and across time. Thus, the MIMIC model is not a standard (linear) simultaneous equation model, but a model which incorporates errors-in-variables, and associated correlations (see Wansbeek and Meijer 2000; Jöreskog and Goldberger 1975).[32] The MIMIC model (and the longitudinal SEM) can also produce an across time "index" of successful settlement which informs about the dynamics of successful settlement (see, however, the previous discussion of "coarse" factor score estimation)—but such an index does not come out of the reduced form specification.[33]

In summary, although somewhat complex (partially as methods used may not be familiar to economists interested in outcomes for immigrants), MIMIC models provide an analytical method that appropriately deals with many important issues that are not dealt with by less sophisticated methods.

Limitations of the Empirical Models

There are a number of limitations of the analysis in this paper:

- The selected MIMIC models are congruent with the data, and plausible, but they are one set of many potential plausible MIMIC models that have not been tested. Nonetheless, the selected models are informative about the immigrant settlement process.
- The MIMIC model allows path coefficients (i.e. factor loadings) to vary across reflective indicators, but inherently imposes the restriction that the loading for each indicator does not vary as a function of the causal variables (e.g. the loadings are the same for economic immigrants irrespective of whether they are from an English speaking background country or not). An avenue for further examination is, in the event of a larger data set becoming available, examination of MIMIC models at a finer level

[31] For example, where y represents reflective indicators (with λ path coefficients), and x formative indicators (with γ path coefficients): $y_t = \Pi x_t + v_t$ where $\Pi = \lambda\gamma'$ and the distribution of v is a function of the variance of the measurement errors. Thus, individual path coefficients (λs and γs)—the parameters of particular interest—can only be extracted if the variance and covariance of measurement errors in formative and reflective indicators are zero (Goldberger 1972; Jöreskog and Goldberger 1975; Wansbeek and Meijer 2000)—an untenable assumption for much of the LSIA data used in this paper.

[32] The exchange between Breusch (2005) and Dell'Anno and Schneider (2006) and Tedds and Giles (2005), regarding estimates of the shadow economy by MIMIC models, is very informative.

[33] On the other hand, if desired, the reduced form can be extracted from the MIMIC specification.

of disaggregation. Similarly, models with alternative parameterisations (e.g. constraining coefficients on formative indicators to be equal across time) may be estimated with increased observations.

- All variables are assumed to have a linear relationship with successful settlement—restricted partly by the time period of analysis (e.g. the period is not long enough to allow diminishing or increasing duration (and age) impacts[34]).
- Qualitative information would be useful to confirm, or explain, some unexpected results (e.g. speculation that immigrant expectations are responsible for the negative association between the PDI and SucSet).

In summary, models could be improved if data were collected specifically to conduct MIMIC models, and the information provided would be enhanced if further waves of LSIA1 and LSIA2 were collected.

Conclusion

The MIMIC models, particularly when assessed for the more recent LSIA2 immigrants, suggest that the current selection procedures aimed at job-ready skilled immigrants do contribute to immigrant successful settlement, but labour market success is only one of several factors that influence immigrant successful settlement, and it is not the most influential factor—with less impact for non-economic than economic immigrants (and clearly no influence for those who are non-participants).

The settlement experience of immigrants can be influenced; and the MIMIC models suggest ways that the settlement process can be improved for all immigrants—those in the labour force as well as previously ignored NLF immigrants.

The settlement path for all immigrants is important in its own right—as the well-being of members of the population should be an important concern of policy makers (Shields and Wheatley Price 2005). In addition, it has a further importance for NLF immigrants; most NLF immigrants are the partner of another—immigration is not necessarily a solitary undertaking (Cobb-Clark and Connolly 2001). Thus, immigration and remigration decisions are generally a family decision. Poorly settled NLF immigrants may influence a highly skilled, quality, immigrant to leave Australia for family reasons—and the family may include younger dependents who would have made a valuable contribution to Australia's population (and age distribution). If an emigration decision is related to the "score" for SucSet, the MIMIC models

[34] Note that including Age as linear or logarithmic did not alter the results: and that all continuous variables are rescaled but remain linear.

suggest immigrants are more likely to leave some time after wave 2 interviews, which may be closer to the point at which they begin to repay Australia's investment in their settlement. This suggests a role for immigrant support policies that extend beyond the first 12 months after arrival.

There is strong evidence of the usefulness of a multidimensional construct for successful settlement, but there is little evidence of Australian government interest in such measures as aids to policy evaluation and policy formulation (Cummins *et al.* 2003; Shields and Wooden 2003a). If however, appropriate policy is to be formulated, successful settlement should not simply be equated with immigrant employment. While labour market success is important to immigrants who are in the labour force (and probably for their partner), it is just one aspect of successful settlement. Addressing immigrant successful settlement more broadly also recognises that many immigrants do not enter the labour force,[35] and hence their settlement progress has, to date, generally been ignored.

The hypothesised and tested MIMIC model of immigrant successful settlement in this paper demonstrate that a number of immigrant characteristics (i.e. formative indicators) are statistically significantly associated with greater or lesser settlement success—and that the impact of these characteristics differs for the three groups of immigrants considered (economic immigrants, non-economic immigrants who are labour force participants, and NLF immigrants). Models range from reasonably straightforward (seven significant formative indicators for NLF immigrants in LSIA2) to quite complex (16 for NLF immigrants in LSIA1). LSIA2 models tend to be less complex, but are able to explain less of the variance in SucSet—perhaps LSIA1 models can be considered superior because three waves of data are more informative—they are more able to capture the progress of the settlement experience. Thus, without wave 3, there would be a tendency to assume SucSet increased with time—but this is not the message from LSIA1. The failure to conduct a wave 3 survey for LSIA2 restricts comparisons over the medium-term, and failure to follow LSIA1 immigrants beyond wave 3 restricts assessment of the longer-term.

The findings in this paper extend the understanding of the settlement process. A refined picture about the role of multiple causal variables is provided—some things matter, but not for all immigrant groups, and others do not appear to be particularly relevant. For example, labour market outcome (measured as the LMSI) is statistically significant at some point in the

[35] That is, in LSIA2, almost a third of offshore visaed immigrant had not entered the labour force after 18 months in Australia and in LSIA1, over a quarter had not entered the labour force after about 3½ years.

settlement process for all immigrants in the labour market—but with varying degree. Thus, in all cases except non-economic immigrants in LSIA1, the impact at wave 2 of the LMSI was between a third and a quarter of the impact at wave 1, and in LSIA1 the labour market did not influence SucSet for economic immigrants by wave 3 (although, as with all significant formative indicators, previous labour market success indirectly influences later SucSet through the process by which current SucSet is influenced by previous SucSet).

An advantage of the MIMIC approach is that it is not necessary to rely on exact measures—each reflective indicator in the measurement model represents a noisy signal of SucSet, recognising that there is no single variable called SucSet (Skrondal and Rabe-Hasketh 2004). Results of the MIMIC models for economic, non-economic, and NLF immigrants in the LSIA show that unobservable SucSet can be represented in models by four easily obtained observable reflective indicators. The measurement model uncovers differences between the three groups of immigrants considered, and the structural (causal) model confirms that SucSet follows a time-path. Thus, the MIMIC model leads to a more comprehensive understanding of the process of immigrant settlement.

Finally, the results of models in this paper support the view that for immigrants in the labour force, labour market success is only one of several factors influencing successful settlement. For immigrants who do not enter the labour force, successful settlement can also be investigated and quantified—these immigrants need not be ignored. Successful settlement can be measured, and modelled, and hence the path to settlement success can be influenced by appropriate government intervention. If, future labour force and population requirements require a greater emphasis on successful settlement (as well as labour market success) the methods in this paper show how the progress of settlement can be tracked, and if necessary, influenced.

References

Bijleveld, C. C. J. H., L. J. T. van der Kamp, et al. (1998). Structural Equation Models for Longitudinal Data. *Longitudinal Data Analysis Designs, Models and Methods*. C. C. J. H. Bijleveld, L. J. T. van der Kamp, A. Mooijaartet al. London, Sage Publications**:** 425.

Bjorner, J. B. & T. S. Kristensen (1999). "Multi-Item Scales for Measuring Global Self-Rated Health: Investigation of Construct Validity Using Structural Equation Models." *Research on Aging* **21**(3): 417-439.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, John Wiley & Sons, Inc.

Boniface, D. R. & M. E. Tefft (1997). "The Application of Structural Equation Modelling to the Construction of an Index for the Measurement of Health-Related Behaviours." *The Statistician* **46**(4): 505-514.
Boomsma, A. (2000). "Reporting Analyses of Covariance Structures." *Structural Equation Modeling* **7**(3): 461-483.
Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, The Guilford Press.
Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming*. Mahwah, New Jersey, Lawrence Erlbaum Associates.
Chou, C. P., P. M. Bentler, et al. (2000). A Two-Stage Approach to Multilevel Structural Equation Models: Publication to Longitudinal Data. *Modeling Longitudinal and Multilevel Data*. T. D. Little, K. U. Schnabel and J. Baumert. Mahwah, New Jersey, Lawrence Erlbaum Associates.
Chung, M. C., I. Dennis, et al. (2005). "A Multiple-Indicator Multiple-Cause Model for Posttraumatic Stress Reactions: Personality, Coping, and Maladjustment." *Psychosomatic Medicine* **67**: 251-259.
Cobb-Clark, D. A. (2001). "The Longitudinal Survey of Immigration to Australia." *The Australian Economic Review* **34**(4): 467-477.
Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ, Lawrence Erlbaum Associates.
Cummins, R. A., R. Eckersley, et al. (2003). "Developing a National Index of Subjective Wellbeing: The Australian Unity Wellbeing Index." *Social Indicators Research* **64**: 159-190.
De Leeuw, J., C. C. J. H. Bijleveld, et al. (1996). The Mixing Approach as a Unifying Framework for Dynamic Multivariate Analysis, CiteSeer.IST.
De Leeuw, J., C. C. J. H. Bijleveld, et al. (1997). "Latent Variables, State Spaces, and Mixing." *Kwantitatieve Methoden* **55**: 95-111.
DIMIA (2002). User Documentation: Longitudinal Survey of Immigrants to Australia. Canberra, Department of Immigration and Multicultural and Indigenous Affairs.
Fleishman, J., W. D. Spector, et al. (2002). "Impact of Differential Item Functioning on Age and Gender Differences in Functional Disability." *Journal of Gerontology* **57B**(5): S275-S284.
Graff, J. & P. Schmidt (1985). Structural Equation Models with Qualitative Observed Variables. *Measuring the Unmeasurable*. P. Nijkamp, H. Leitner and N. Wrigley. Dordrecht, Martinus Nijhoff Publishers.
Hayduk, L. A. (1996). *LISREL Issues, Debates, and Strategies*. Baltimore, The John Hopkins University Press.
Hellgren, J. & M. Sverke (2003). "Does Job Insecurity Lead to Impaired Well-being or Vice Versa? Estimation of Cross-Lagged Effects Using Latent Variable Modelling." *Journal of Organizational Behavior* **24**: 215-236.
Jöreskog, K. G. (2004). Structural Equation Modeling with Ordinal Variables Using LISREL (Revised 2006). Chicago, Scientific Software International, Inc.
Jöreskog, K. G. & D. Sörbom (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, Scientific Software International, Inc.
Jöreskog, K. G. & D. Sörbom (2002). *PRELIS 2: User's Reference Guide*. Lincolnwood, IL, Scientific Software International, Inc.
Kaplan, D. (2000). *Structural Equation Modeling: Foundations and Extensions*. Thousand Oaks, Sage Publications, Inc.
Kenny, D. A. (1979). *Correlation and Causality*. New York, Wiley.

Kim, H. & J. W. Rojewski (2002). "Using Structural Equation Modeling to Improve Research in Career and Technical Education." *Journal of Vocational Education Research* **27**(2): 257-274.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. New York, Guilford Press.

Kline, R. B. (2006). Reverse Arrow Dynamics. *Structural Equation Modeling: A Second Course*. G. R. Hancock and R. O. Mueller. Greenwich, CT, Information Age Publishing**:** 43-68.

Lei, P.-W. & Q. Wu (2007). "Introduction to Structural Equation Modeling: Issues and Practical Considerations." *Instructional Topics in Educational Measurement* **Fall**: 33-43.

Lester, L. H. (2008). Measuring, Modelling, and Monitoring the Dynamics of Labour Market Success and Successful Settlement of Immigrants to Australia, National Institute of Labour Studies, Adelaide, Flinders University, Unpublished Doctor of Philosophy.

Lester, L. H. (2007). An Index of Labour Market Success. *Working Paper 156*. Adelaide, National Institute of Labour Studies.

Long, S. J. (1983). *Confirmatory Factor Analysis: A Preface to LISREL*. Beverly Hills, Sage Publications.

Markowitz, F. E. (2001). "Modeling Processes in Recovery from Mental Illness: Relationships Between Symptoms, Life Satisfaction, and Self-Concept." *Journal of Health and Social Behavior* **42**(1): 64-79.

Marsh, H. W., D. K. Tracey, et al. (2006). "Multidimensional Self-Concept Structure for Preadolescents With Mild Intellectual Disabilities: A Hybrid Multigroup-MIMIC Approach to Factorial Invariance and Latent Mean Differences." *Educational and Psychological Measurement* **66**(5): 795-818.

Masterson, S., K. Lewis, et al. (2000). "Integrating Justice and Social Exchange: The Differing Effects of Fair Procedures and Treatment on Work Relationships." *Academy of Management Journal* **43**(4): 738-748.

Meade, A. W. & D. J. Bauer (2007). "Power and Precision in Confirmatory Factor Analytic Tests of Measurement Invariance." *Structural Equation Modeling* **14**(4): 611-635.

Montfort, K. V. & C. C. J. H. Bijleveld (2004). "Dynamic Analysis of Multivariate Panel Data with Non-Linear Transformations." *Journal of Mathematical Psychology*(48): 322-333.

Muthén, B. O. (1993). Goodness of Fit with Categorical Data and Other Nonnormal Variables. *Testing Structural Equation Models*. K. A. Bollen and J. S. Long. Newbury Park, CA, Sage.

Pardo, R. & F. Calvo (2002). "Attitudes Towards Science Among the European Public: A Methodological Analysis." *Public Understanding of Science* **11**: 155-195.

Richardson, S., L. Miller-Lewis, et al. (2002). The Settlement Experiences of New Migrants: A Comparison of Wave One of LSIA 1 and LSIA 2. *Report to the Department of Immigration and Multicultural and Indigenous Affairs*. Canberra, DIMIA.

Schumacker, R. E. & R. G. Lomax (1996). *A Beginner's Guide to Structural Equation Modeling*. Mahwah, New Jersey, Lawrence Erlbaum Associates.

Shields, M. A. & M. Wooden (2003). Investigating the Role of Neighbourhood Characteristics in Determining Life Satisfaction. *Working Paper No. 24/03*. Melbourne, Melbourne Institute of Applied Economic and Social Research.