

Instrumental-Variable Estimation of Fixed-Effects Fractional-Outcome Panel Data Models, with an Application to a Major Disability Support Policy Reform*

Abstract

We develop a fixed-effects estimation framework for fractional-outcome panel models with a logit conditional mean. Pairwise quasi-differencing eliminates individual effects from the moment conditions, yielding a generalized method of moments estimator that is consistent, under the stated moment restrictions, even when the number of time periods is finite and fixed. The framework allows unrestricted dependence between individual effects and covariates, accommodates endogenous regressors through instrumental variables, and extends to multivariate shares that sum to one. Boundary observations require no adjustment, and average semi-elasticities can be estimated without recovering the individual effects. Monte Carlo experiments show little bias in the fractional-outcome designs considered and precision gains from combining additional valid time pairs. We apply the framework to Australia's National Disability Insurance Scheme (NDIS), using HILDA (Household, Income and Labour Dynamics in Australia) survey data for 2011–2023 and instruments based on the scheme's staggered geographic rollout. We examine how NDIS receipt affects the allocation of time across paid work, housework, disability care, and remaining activities for recipients and other household members. The instrumental-variable estimates indicate that recipients reduce time in paid work and in providing disability care, while other household members increase disability care and reduce time devoted to remaining activities.

JEL classification

C23, C24, I18, J22

Keywords

panel data, fixed effects, fractional responses, share data, time use, disability support

Corresponding author

Kevin E. Staub

kevin.staub@unimelb.edu.au

* *Acknowledgments:* We are thankful to Myoung-Jae Lee, Rainer Winkelmann, Mark Harris, Chris Rose, Otavio Bartalotti, as well as seminar participants at the University of Queensland, the University of Lucerne, the University of Melbourne, the Conference of the International Association for Applied Econometrics in Turin, the International Panel Data Conference in Montpellier, the European Winter Meeting of the Econometric Society in Nicosia, the Australasian Workshop on Econometrics and Health Economics in Port Stephens, and the Conference for Applied Econometrics in Montpellier for helpful comments on earlier versions.

1 Introduction

Empirical questions about the allocation of time, expenditure, and other limited resources naturally lead to models for fractional outcomes.¹ These outcomes are bounded between zero and one, often include observations at the boundaries, and may form a system of shares that sum to one. Fractional response models provide a convenient way to respect these features while allowing covariate effects to vary with the conditional mean. In panel applications, however, researchers must also account for persistent unobserved differences across individuals or firms and, frequently, endogenous explanatory variables. Preferences, productive capacity, or support needs may affect both an observed allocation and the decisions or policies that determine it. Adding individual intercepts to a nonlinear model does not generally resolve this problem when the panel is short, because estimation of the common parameters is subject to the incidental parameter problem. Applied researchers therefore need methods that accommodate these sources of endogeneity while preserving the structure of fractional outcomes.

This paper develops a fixed-effects estimation framework for fractional panel data models with a logit conditional mean. Its central feature is a pairwise quasi-differencing transformation that eliminates the individual effects from the moment conditions.² The resulting generalized method of moments (GMM) estimator is consistent as the number of individuals increases with the number of periods fixed, subject to identification and the stated moment restrictions. Estimation does not require specifying a distribution for the individual effects or restricting their association with the explanatory variables. The framework also accommodates instrumental variables and extends to multivariate fractional outcomes with a multinomial logit conditional mean. Consequently, it provides a common approach to estimating a single fraction or a system of shares in short panels with correlated individual heterogeneity and potentially endogenous regressors.

The approach builds on the conditional-mean formulation of fractional regression developed by [Papke and Wooldridge \(1996\)](#). Their quasi-likelihood method accommodates fractional outcomes, including zeros and ones, without specifying their full conditional distribution. For a survey of cross-sectional fractional response models, see [Ramalho, Ramalho and Murteira \(2011\)](#). Existing panel methods address unobserved heterogeneity through several routes. The correlated random-effects approach of [Papke and Wooldridge \(2008\)](#) models the relationship between individual heterogeneity and the covariates and can be combined with a control-function treatment of endogeneity (see also [Bates, Papke and Wooldridge, 2024](#)). Binomial fixed-effects methods, examined by [Machado \(2004\)](#) and [Winkelmann and Xu \(2021\)](#), exploit the structure of outcomes represented as proportions of a fixed

¹Recent applications include contributions to the economics of health ([Mullahy, 2024](#); [Mullahy and Robert, 2010](#)), labor ([Bedard and Deschênes, 2006](#)), education ([Bates, Papke and Wooldridge, 2024](#)), international trade ([Santos Silva, Tenreiro and Wei, 2014](#)), and industrial organization ([Azoulay, Waverly and Stuart, 2009](#)), among many others.

²[Bonhomme \(2012\)](#) provides a general functional-differencing framework for models with a parametrically specified outcome distribution conditional on covariates and individual effects. Here, we specify the logistic conditional mean and cross-period moment restrictions, leaving other distributional features unspecified. With infinite supports for outcomes and individual effects, as permitted in our framework, the space of admissible moment functions may be infinite-dimensional, and their selection can affect finite-sample precision.

number of trials. [Ramalho, Ramalho and Coelho \(2018\)](#) develop fixed-effects estimators using an exponential regression representation and restrictions on the transformed model. Our contribution is to derive fixed-effects moment conditions for the logistic conditional mean itself, applicable to general fractional outcomes and suitable for instrumental-variable estimation. This preserves a familiar specification while avoiding estimation of the individual intercepts.

The moment conditions combine outcomes from two periods with an exponential function of the change in covariates. Under the logit mean specification, this combination removes the individual effect. Validity also requires restrictions on the errors across the periods being paired: in the baseline model, they have zero conditional means and zero conditional covariance given the covariates for the pair and the individual effect. These restrictions need only hold for the pairs used in estimation. The choice of time pairs can therefore accommodate some departures from strict exogeneity over the full regressor history; we demonstrate this possibility with a particular predetermined-regressor process. When regressors are endogenous, excluded instruments can instead be used under orthogonality conditions for the quasi-differenced residual. Expressing identification through these transformed moment conditions makes the assumptions underlying the instrumental-variable estimator explicit and permits estimation within the standard GMM framework.

The multivariate extension is useful when the economic question concerns the allocation of a common total across competing uses. Modeling time or expenditure shares jointly imposes adding-up restrictions and allows a covariate to affect the entire allocation. We use the multinomial logit conditional-mean specification employed in multivariate fractional regression by [Mullahy \(2015\)](#), with outcome-specific individual effects. Pairwise transformations eliminate these effects and generate moment conditions for the slope parameters of each share equation. The formulation allows dependence between the shares within a period, while imposing the relevant cross-period restrictions for the selected pairs. The same construction accommodates instrumental variables, allowing endogenous determinants of an allocation to be analyzed within a coherent system of fractional outcomes.

Several features make the framework convenient for empirical work. The transformation is defined at both zero and one, so boundary observations require no adjustment. Binary and multinomial discrete outcomes are included as special cases, allowing related participation outcomes and fractional allocations to be studied using a common estimation approach. Although the individual effects are eliminated, economically interpretable summaries remain available: average semi-elasticities can be estimated from the slope coefficients and observed outcome averages, both for the population and for subgroups. Their estimation does not require recovering the individual effects. Moreover, the dimension of the parameter vector to be estimated does not increase with the number of individuals. These features simplify implementation and interpretation in applications with many individuals observed over relatively few periods.

We use simulation experiments to examine the finite-sample performance of the estimators for univari-

ate and multivariate outcomes. The designs vary the length of the panel, the prevalence of boundary observations, and the relationship between individual heterogeneity and the regressors, and include the predetermined-regressor example. The proposed estimators exhibit little bias in the fractional-outcome designs considered. The comparisons illustrate the consequences of ignoring correlated individual heterogeneity, estimating individual intercepts in short nonlinear panels, or using a transformed-outcome specification that does not match the data-generating process. Combining additional valid time pairs generally improves precision. The experiments also document efficiency trade-offs relative to estimators that exploit more specific features of the outcome distribution, providing guidance on the practical scope of the proposed approach.

We apply the framework to study how Australian households changed their time allocation across various categories of paid and unpaid work after receiving support from the new National Disability Insurance Scheme (NDIS). The NDIS, introduced from 2013–2020, embodies a market-based approach to care, allowing people with a disability to have control over the type of government-funded support they receive and who provides that support. The NDIS is estimated to have cost AUD 48.5 billion in the 2024–2025 financial year, making it Australia’s second largest social security and welfare program ([Australian Government, 2025](#)). The scheme has been criticized for exceeding original budget forecasts; accordingly, a rigorous assessment of the impacts of the scheme on participants is critical. Our objective here is to contribute new causal evidence towards this by investigating a multivariate fractional-outcome model of time use that models the shares of weekly time devoted to paid work, housework, disability care, and all remaining activities. One of the aims of the NDIS is to improve labor market outcomes for recipients, with potential economic gains also anticipated for family members ([Productivity Commission, 2011](#)). Funded services may remove barriers to paid work and substitute for informal care, but they may also reduce the need to work to finance support or create additional demands for coordinating services. The implications for time allocation can consequently differ between NDIS recipients and other household members. To date, there are very few quantitative studies that estimate the impact of the NDIS on labor elasticity ([Ma et al., 2025](#); [Tan, 2025](#)), and none that look at it through the more detailed and comprehensive lens of households’ time allocation.

We obtain information on time use and NDIS receipt from the Household, Income and Labour Dynamics in Australia (HILDA) survey, covering the years 2011–2023 in our baseline analysis. Separate indicators for a respondent’s own NDIS receipt and receipt by another household member allow us to examine both direct effects and effects within the household. Accounting for fixed effects is essential in this setting to control for unobserved, time-invariant preferences, skills and constraints in abilities, which might be correlated with the household obtaining access to NDIS support. However, NDIS receipt might be further correlated with time-varying unobservables, complicating the estimation of causal effects. Consequently, we propose a novel instrumental variable (IV) strategy, exploiting exogenous variation in NDIS receipt generated by the step-wise geographic roll-out of the program over time. This roll-out implied that in a given year, some of the individuals who in principle would meet

the eligibility criteria for NDIS receipt were ineligible due to the NDIS not (yet) being available in their area of residence. For this, we supplement the HILDA data with information on the precise location and dates of the NDIS eligibility roll-out compiled from administrative government sources.

The instrumental-variable estimates indicate that NDIS receipt reduces recipients' time in paid work and in providing disability care. For other household members, the estimates indicate an increase in disability care and a reduction in time available for remaining activities. The response of household members' care time also varies with the recipient's age: the estimates indicate a reduction when recipients are younger than 15 and an increase when recipients are aged 15 or over. Complementary binary-outcome estimates suggest that the extensive margin effect of changes in participation plays an important role in the responses in paid work and care.

Section 2 develops the univariate and multivariate estimators, discusses instruments and interpretation, and presents the Monte Carlo evidence. Section 3 describes the NDIS application and reports the empirical results. Section 4 concludes.

2 Econometric framework

Consider the estimation of a regression function of a bounded scalar dependent variable $0 \leq y \leq 1$, say, conditional on a vector of K explanatory variables x . Let the $(k \times 1)$ vector x_{it} denote a set of observations on the variables x for individual i at time t , with $i \in \{1, 2, \dots, n\}$ and $t \in \{1, \dots, T\}$. Suppose that the regression function of interest is

$$\mathbb{E}(y_{it}|x_{it}, \alpha_i) = \frac{\exp(\alpha_i + x'_{it}\beta)}{1 + \exp(\alpha_i + x'_{it}\beta)} \equiv \Lambda(\alpha_i + x'_{it}\beta), \quad (1)$$

where β is a K -vector of coefficients and the scalar α_i is an individual-specific fixed effect. In what follows we assume that N is large while T is small and fixed. Even in the special case when we assume that $\alpha_i = 0$ for all individuals, the bounded nature of the dependent variable and the possibility of nontrivial probability mass at the boundaries raise some estimation and inference issues. A solution to these issues in this special case is the logit PL approach of Papke and Wooldridge (1996), which consists in estimating the parameters β by maximizing the (pseudo-)log-likelihood; that is, estimating β by

$$\hat{\beta} = \arg \max_{\beta} \log \ell_n(\beta) = \arg \max_{\beta} \left[\sum_i \sum_t y_{it} \log \Lambda(x'_{it}\beta) + (1 - y_{it}) \log(1 - \Lambda(x'_{it}\beta)) \right].$$

However, a key motivation for using panel data is to be able to account for individual-specific effects α_i when estimating β . One possibility is to estimate the α_i directly, for instance by including a set of individual dummy variables (Hausman and Leonard, 1997; Wagner, 2001), but this is subject to the well-known incidental parameter problem (Neyman and Scott, 1948) and leads to inconsistent

estimates of β (Machado, 2004). Another possibility is to specify a distribution for α_i , such as Papke and Wooldridge's (2008) correlated random effects approach, in which the α_i are assumed to be normally distributed. Instead of placing restrictions on the distribution of α_i , an alternative approach is to leave them unrestricted and instead place restrictions on the distribution of y_{it} , such as requiring it to be binomial (Winkelmann and Xu, 2021; Machado, 2004). Finally, one can abandon the logistic conditional mean function (1). For instance, Ramalho, Ramalho and Coelho (2018) provide a fixed effects estimator by specifying $y_{it} = \Lambda(\alpha_i + x'_{it}\beta + \varepsilon_{it})$, where ε_{it} is an IID idiosyncratic error. Thus, marginal of ε_{it} , the conditional mean $E(y_{it}|x_{it}, \alpha_i)$ is not logistic.

We choose to stay within the framework of a logit fixed effects model. One advantage of retaining the logit functional form is the comparability it ensures with other logit-based models such as, for instance, the binary logit or the multinomial logit. A second advantage is that each coefficient β_k is proportional to the corresponding average semi-elasticity, which can be estimated easily, as shown by Kitazawa (2012) and Kemp and Santos Silva (2016) for binary response variables. Specifically, from the conditional mean function, it follows that the semi-elasticity for individual i at time t with respect to the k th explanatory variable $x_{k,it}$ is

$$e_{it,k} = \frac{\partial \log \mathbb{E}(y_{it}|x_{it}, \alpha_i)}{\partial x_{k,it}} = \beta_k \frac{1}{1 + \exp(x'_{it}\beta + \alpha_i)},$$

where β_k denotes the k th element in β , and the average semi-elasticity $\mathbb{E}(e_{it,k}) = \beta_k(1 - \mathbb{E}(y_{it}))$ can be consistently estimated as

$$\widehat{\mathbb{E}}(e_{it,k}) = \widehat{\beta}_k(1 - \bar{y}), \quad \bar{y} = N^{-1}T^{-1} \sum_i \sum_t y_{it},$$

provided $\widehat{\beta}$ is a consistent estimator of β .³ We show below how to obtain such a large- N consistent estimator within the framework of the logit fixed effects model.

2.1 Quasi-differenced fractional logit estimator

While our proposed estimator allows for an arbitrary number of time periods and can be applied to a multivariate dependent variable, to fix ideas we first present a version of it for the univariate case of a scalar y_{it} .

Construction of moment conditions. We rewrite the conditional mean model in error form as

$$y_{it} = \Lambda_{it} + \epsilon_{it}, \quad i = 1, \dots, N; t = 1, \dots, T,$$

³Conditional average semi-elasticities for subgroups can be obtained analogously. For a subgroup defined by $x_{it,l} = x_{0,l}$, such as women or employed individuals, $\mathbb{E}(e_{it,k} | x_{it,l} = x_{0,l}) = \beta_k[1 - \mathbb{E}(y_{it} | x_{it,l} = x_{0,l})]$. This quantity can be estimated by $\widehat{\beta}_k(1 - \bar{y}_{x_l=x_{0,l}})$, where $\bar{y}_{x_l=x_{0,l}}$ denotes the sample mean of y_{it} over observations belonging to that subgroup.

where we define $\Lambda_{it} \equiv \Lambda(\alpha_i + x'_{it}\beta)$, and ϵ_{it} is the conditional-mean residual. Since $y_{it} \in [0, 1]$, this residual is bounded between $-\Lambda_{it}$ and $1 - \Lambda_{it}$ and will generally be conditionally heteroskedastic. Let $\mathcal{P} \subseteq \{(t, s) : 1 \leq s < t \leq T\}$ denote a set of time pairs used in estimation. For every $(t, s) \in \mathcal{P}$, we impose

$$\mathbb{E}(\epsilon_{ir} \mid x_{it}, x_{is}, \alpha_i) = 0, \quad r \in \{t, s\}, \quad (2)$$

$$\mathbb{E}(\epsilon_{it}\epsilon_{is} \mid x_{it}, x_{is}, \alpha_i) = 0. \quad (3)$$

The first condition, (2), is a pair-specific form of strict mean exogeneity: the conditional-mean residual in either period is mean independent of the explanatory variables in both periods used to construct estimation moments. The second condition, (3), requires conditional serial uncorrelatedness across those two periods. Conditional independence of y_{it} and y_{is} given $(x_{it}, x_{is}, \alpha_i)$ is sufficient for condition (3), but is not necessary.

Importantly, conditions (2)–(3) need hold only for the time pairs included in $\mathcal{P}$. The stronger assumptions of strict exogeneity with respect to the complete covariate history and conditional serial independence would imply these conditions for every pair, but are not required. For example, if dependence is confined to adjacent periods while conditions (2)–(3) hold for periods separated by at least two time intervals, the estimator can be constructed from the more distant time contrasts while excluding adjacent-period contrasts.

Consider the pairwise products of y_{it} and $1 - y_{is}$ for some $(t, s) \in \mathcal{P}$.⁴

$$y_{it}(1 - y_{is}) = \Lambda_{it}(1 - \Lambda_{is}) + \epsilon_{it}(1 - \Lambda_{is}) - \epsilon_{is}\Lambda_{it} - \epsilon_{it}\epsilon_{is}, \quad (4)$$

$$y_{is}(1 - y_{it}) = \Lambda_{is}(1 - \Lambda_{it}) + \epsilon_{is}(1 - \Lambda_{it}) - \epsilon_{it}\Lambda_{is} - \epsilon_{it}\epsilon_{is}. \quad (5)$$

Observe that

$$\frac{\Lambda_{it}(1 - \Lambda_{is})}{\Lambda_{is}(1 - \Lambda_{it})} = \exp((x'_{it} - x'_{is})\beta)$$

is free from α_i . Combining this expression with (4) and (5), we obtain

$$\begin{aligned} & y_{it}(1 - y_{is}) - \exp((x'_{it} - x'_{is})\beta)y_{is}(1 - y_{it}) \\ &= \epsilon_{it}(1 - \Lambda_{is}) - \epsilon_{is}\Lambda_{it} - \epsilon_{it}\epsilon_{is} - \frac{\Lambda_{it}(1 - \Lambda_{is})}{\Lambda_{is}(1 - \Lambda_{it})} [\epsilon_{is}(1 - \Lambda_{it}) - \epsilon_{it}\Lambda_{is} - \epsilon_{it}\epsilon_{is}]. \end{aligned} \quad (6)$$

After taking expectations, given our assumptions, we arrive at the moment condition

$$\mathbb{E}[y_{it}(1 - y_{is}) - \exp((x'_{it} - x'_{is})\beta)y_{is}(1 - y_{it}) \mid x_{it}, x_{is}, \alpha_i] = 0, \quad (7)$$

Writing $u_{i,t,s}(\beta) = y_{it}(1 - y_{is}) - \exp(\Delta x'_{i,t,s}\beta)y_{is}(1 - y_{it})$ and $\Delta x_{i,t,s} = x_{it} - x_{is}$, we can rewrite the

⁴The estimator works even with $T = 2$, which is when the incidental parameters bias is at its most severe. If $T = 2$ then we have, for example, $s = 1$ and $t = 2$.

conditional moment compactly as

$$\mathbb{E}[u_{i,t,s}(\beta) \mid x_{it}, x_{is}] = 0. \quad (8)$$

Consequently, for any $Q \times 1$ vector of functions $q_{i,t,s} = q(x_{it}, x_{is})$, with $Q \geq K$, we obtain the unconditional moments

$$\mathbb{E}[u_{i,t,s}(\beta)q_{i,t,s}] = 0. \quad (9)$$

In particular, we initially consider $q_{i,t,s} = \Delta x_{i,t,s}$. In this case, the instrument vector contains $Q = K$ elements. Since the unconditional moments do not depend on the incidental parameters α_i , we can use them as a basis for obtaining a consistent GMM estimator for β .

GMM estimator. To construct a GMM estimator from (9), two specification choices have to be made. First, which t, s to use. For instance, one may choose to use only $s = t - 1$, which results in $T - 1$ pairwise moment contributions per individual. We call this the ‘first differences’ form of the estimator. If using all possible differences, there are $\binom{T}{2} = T(T - 1)/2$ pairwise moment contributions per individual. Second, if using higher-order differences, in addition to first differences, one might want to treat each order of difference as a distinct set of moment conditions, which leads to overidentification, rather than pooling the orders as additional pairwise contributions. We expand on these issues and give further examples below.

A general representation of the GMM estimator for β that encompasses both forms can be written as

$$\hat{\beta} = \arg \min_{\beta} \bar{h}_N(\beta)' \mathbf{W} \bar{h}_N(\beta), \quad (10)$$

where

$$\bar{h}_N(\beta) = \frac{1}{N} \sum_{i=1}^N h_i(\beta).$$

Here $\mathbf{W}$ is a conformable positive semidefinite weighting matrix and $h_i(\beta)$ is an individual-level moment function. To define possible forms of $h_i(\beta)$, write

$$m_{i,t,s}(\beta) = u_{i,t,s}(\beta)q_{i,t,s}.$$

Let $\mathcal{R} \subseteq \{1, \dots, T - 1\}$ denote the set of included difference orders. For each $r \in \mathcal{R}$, define

$$\mathcal{P}_r = \{(t, t - r) : t = r + 1, \dots, T\}, \quad M_r = |\mathcal{P}_r| = T - r, \quad M = \sum_{r \in \mathcal{R}} M_r.$$

The chosen orders are restricted so that $\mathcal{P}_r \subseteq \mathcal{P}$ for every $r \in \mathcal{R}$, where $\mathcal{P}$ is the set of time pairs for which the moment conditions are valid.

If different difference orders are treated as additional observations, their moment contributions are

pooled and the individual-level moment function is

$$h_i^P(\beta) = \frac{1}{M} \sum_{r \in \mathcal{R}} \sum_{t=r+1}^T m_{i,t,t-r}(\beta). \quad (11)$$

This moment function is of dimension $Q \times 1$, and $\mathbf{W}$ is therefore $Q \times Q$. Alternatively, if the moment contributions for each order are treated as a distinct set of moment conditions, the individual-level moment function is stacked as

$$h_i^S(\beta) = \begin{pmatrix} \frac{1}{M_{r_1}} \sum_{t=r_1+1}^T m_{i,t,t-r_1}(\beta) \\ \vdots \\ \frac{1}{M_{r_R}} \sum_{t=r_R+1}^T m_{i,t,t-r_R}(\beta) \end{pmatrix}, \quad \mathcal{R} = \{r_1, \dots, r_R\}. \quad (12)$$

This moment function is of dimension $RQ \times 1$, and $\mathbf{W}$ is therefore $RQ \times RQ$.

The distinction between pooling and stacking moment restrictions is related to the classic panel-GMM literature, which emphasizes how implied orthogonality conditions are organized for estimation; see (Arellano and Bover, 1995; Ahn and Schmidt, 1995, 1997). Pooling moment contributions across time gaps produces an average restriction, whereas stacking them by time gap treats the restrictions as distinct moments and may generate overidentification. Although additional valid and informative moments can improve asymptotic efficiency, they may worsen finite-sample performance. We therefore treat the choice of time pairs and whether to pool or stack their contributions as part of the empirical specification.⁵

Examples. Let us consider three examples using the initial choice $q_{i,t,s} = \Delta x_{i,t,s}$, for which $Q = K$. First, if the estimator uses ‘first differences’ only, $\mathcal{R} = \{1\}$ and

$$h_i(\beta) = \frac{1}{T-1} \sum_{t=2}^T m_{i,t,t-1}(\beta).$$

The pooled and stacked forms coincide, and there are $N(T-1)$ underlying pairwise moment contributions. The moment function is of dimension $K \times 1$, so the choice of $\mathbf{W}$ is irrelevant (Greene, 2012, pp. 505–6).

Second, suppose the estimator uses ‘first’ and ‘second differences’, so that $\mathcal{R} = \{1, 2\}$. If the ‘second

⁵In our numerical assessment of the small-sample performance of the estimator (Section 2.3), we use both ‘first differences’ and ‘all differences’ specifications. Computational or other considerations may lead the researcher to opt for using fewer orders of difference. Indeed, using first differences is a common empirical strategy implemented as a default in similar estimators that can in principle use all pairwise contrasts within panel units, for instance, in the statistical software STATA’s command for the Arellano and Bond (1991) estimator. In our empirical specification (Section 3), we use all differences up to the fourth lag order as the baseline. Our robustness checks include specifications with different lag orders and show that our results are qualitatively similar to the baseline.

differences' are treated as additional observations, the pooled moment function is

$$h_i^P(\beta) = \frac{1}{2T-3} \left[\sum_{t=2}^T m_{i,t,t-1}(\beta) + \sum_{t=3}^T m_{i,t,t-2}(\beta) \right].$$

There are $N(2T-3)$ underlying pairwise moment contributions, while the moment function remains of dimension $K \times 1$.

Third, the 'second differences' can instead be treated as an additional set of moment conditions, in which case

$$h_i^S(\beta) = \begin{pmatrix} \frac{1}{T-1} \sum_{t=2}^T m_{i,t,t-1}(\beta) \\ \frac{1}{T-2} \sum_{t=3}^T m_{i,t,t-2}(\beta) \end{pmatrix}.$$

The first block uses $N(T-1)$ pairwise moment contributions and the second uses $N(T-2)$. The resulting moment function is of dimension $2K \times 1$, leading to overidentification.

Endogenous regressors. The preceding derivation treats every element of x_{it} as exogenous. Suppose instead that the covariate vector can be partitioned as

$$x_{it} = \left((x_{it}^{\text{exog}})' , (x_{it}^{\text{endog}})' \right)',$$

where x_{it}^{exog} satisfies the pair-specific exogeneity conditions, while x_{it}^{endog} does not. Let $\Delta z_{i,t,s}$ denote a vector of excluded instruments associated with pair (t, s) . In the leading case, these are differences in a time-varying instrument, $\Delta z_{i,t,s} = z_{it} - z_{is}$; more generally, the notation can collect excluded instruments constructed for the particular pair. The instrument vector is then

$$q_{i,t,s} = \left((\Delta x_{i,t,s}^{\text{exog}})' , (\Delta z_{i,t,s})' \right)'.$$

Thus, the endogenous covariates remain in $u_{i,t,s}(\beta)$, while their differences are excluded from $q_{i,t,s}$ and replaced by suitable instruments. Although the conditional moment (8) need no longer hold with respect to the complete covariate vector, the unconditional moment (9) remains valid provided that

$$\mathbb{E}[u_{i,t,s}(\beta) \Delta x_{i,t,s}^{\text{exog}}] = 0 \quad \text{and} \quad \mathbb{E}[u_{i,t,s}(\beta) \Delta z_{i,t,s}] = 0$$

for every $(t, s) \in \mathcal{P}$ used in estimation. Instrument validity is therefore required only for the selected time pairs.

The number of instruments must satisfy the relevant order condition. In the pooled formulation, this requires $Q \geq K$; in the stacked formulation, the total number of moments, RQ , must be at least K . Instrument count alone is not sufficient for identification. The Jacobian of the final moment vector

must have full column rank,

$$\text{rank} \left\{ \mathbb{E} \left[\frac{\partial h_i(\beta)}{\partial \beta'} \right] \right\} = K.$$

This rank condition requires the excluded instruments to provide identifying variation for the endogenous covariates. Identification also requires variation in the corresponding elements of $\Delta x_{i,t,s}$. If $\Delta x_{i,t,s,k} = 0$ for every pairwise contribution in the estimation sample, then β_k is not identified and $x_{it,k}$ cannot be used as a regressor in the quasi-differenced model.

Two typical sources of endogeneity arise in panel-data applications. First, an element of x_{it}^{endog} may be an endogenous choice or participation variable that is jointly determined with the outcome. We illustrate such a case in detail through our empirical application in Section 3, where access to disability support may be endogenous for this reason. We use program eligibility induced by the progressive geographic roll-out of the program as an instrument, exploiting variation in eligibility that was plausibly exogenous to individuals.

Second, x_{it}^{endog} may be a weakly endogenous or predetermined variable. It may be tempting to apply IV strategies from linear panel-data models directly, for example by using a lagged difference such as $\Delta z_{i,t,t-1} = x_{i,t-2} - x_{i,t-3}$ as an instrument for the moment $u_{i,t,t-1}(\beta)$, in analogy with [Arellano and Bond \(1991\)](#). Such approaches are not generally valid in the present setting because $u_{i,t,t-1}(\beta)$ is nonlinear in the outcomes and endogenous regressors. Consequently, conventional orthogonality between a lagged instrument and the CEF errors does not generally imply $\mathbb{E}[\Delta z_{i,t,t-1} u_{i,t,t-1}(\beta)] = 0$.⁶ In some cases, however, an appropriate choice of the pair (t, s) can accommodate a predetermined regressor without excluded instruments. For example, suppose that $x_{it} = \mu_{it} + \varepsilon_{i,t-1}$, where μ_{it} is strictly exogenous and the CEF errors ε_{it} are serially independent. Then x_{it} is endogenous in $u_{i,t,t-1}(\beta)$, because it depends on one of the errors entering that moment, but is exogenous in $u_{i,t,t-2}(\beta)$. The corresponding $\Delta x_{i,t,t-2} = x_{it} - x_{i,t-2}$ can therefore be used in $q_{i,t,t-2}$ to construct a valid moment, subject to the usual rank condition. We present such a case in Section 2.3, where we explore the finite-sample performance of the estimator through Monte Carlo simulations.

Asymptotic distribution. Since the individuals are independent sampling units, the estimator (10) is consistent and asymptotically normally distributed as $N \rightarrow \infty$ under standard regularity conditions (cf. e.g., [Wooldridge, 2010](#)). The asymptotic variance of $\sqrt{N}(\hat{\beta} - \beta)$ has the usual form, $\mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1}$, where

$$\mathbf{A} = \mathbf{G}' \mathbf{W} \mathbf{G}, \quad \mathbf{B} = \mathbf{G}' \mathbf{W} \mathbb{E} [h_i(\beta) h_i(\beta)'] \mathbf{W} \mathbf{G}, \quad \mathbf{G} \equiv \mathbb{E} \left[\frac{\partial h_i(\beta)}{\partial \beta'} \right].$$

⁶In particular, without modifying the moment conditions, the model does not accommodate specifications containing lags of the dependent variable. An error in an earlier outcome propagates into later regressors through the dynamic recursion, so that $\mathbb{E}[u_{i,t,s}(\beta)] \neq 0$ generally for any finite $t - s$, even if the CEF errors are serially independent.

2.2 Multivariate fractional dependent variable

The idea of the quasi-differenced GMM estimator (10) can be extended to accommodate a panel setting where the outcome is a simplex; that is, several shares that sum to unity. In this setting, the conditional mean takes on the multinomial logit form, with separate fixed effects for each of the outcome shares.

The panel data comprise $(y_{it}^{[1]}, \dots, y_{it}^{[J+1]}, x_{it})$, $i = 1, \dots, N$ and $t = 1, \dots, T$, where each outcome share lies in the interval $[0, 1]$ and $\sum_{j=1}^{J+1} y_{it}^{[j]} = 1$. The last, linearly dependent category is used as the base share, leaving the model with J outcome variables. In particular,

$$y_{it}^{[J+1]} = 1 - \sum_{j=1}^J y_{it}^{[j]}.$$

Similar to the univariate case, any of the outcome shares may take the boundary values zero or one with non-trivial masses. This includes the multinomial-response special case in which one outcome equals unity and all others equal zero. Define $\alpha_i = (\alpha_i^{[1]}, \dots, \alpha_i^{[J]})'$. The conditional mean functions are

$$\mathbb{E}[y_{it}^{[j]} \mid x_{it}, \alpha_i] = \frac{\exp(\alpha_i^{[j]} + x'_{it}\beta^{[j]})}{1 + \sum_{\ell=1}^J \exp(\alpha_i^{[\ell]} + x'_{it}\beta^{[\ell]})} \equiv \Lambda_{it}^{[j]}, \quad j = 1, \dots, J,$$

and

$$\mathbb{E}[y_{it}^{[J+1]} \mid x_{it}, \alpha_i] = \frac{1}{1 + \sum_{\ell=1}^J \exp(\alpha_i^{[\ell]} + x'_{it}\beta^{[\ell]})} \equiv \Lambda_{it}^{[J+1]}.$$

Here $\beta^{[j]}$ is a $K \times 1$ parameter vector and $\alpha_i^{[j]}$ is the fixed effect for individual i and outcome j .

Define the conditional-mean residuals as

$$\epsilon_{it}^{[j]} = y_{it}^{[j]} - \Lambda_{it}^{[j]}, \quad j = 1, \dots, J+1.$$

Because both the observed shares and their conditional means add to unity,

$$\sum_{j=1}^{J+1} \epsilon_{it}^{[j]} = 0, \quad \epsilon_{it}^{[J+1]} = - \sum_{j=1}^J \epsilon_{it}^{[j]}.$$

Moreover, each residual satisfies

$$-\Lambda_{it}^{[j]} \leq \epsilon_{it}^{[j]} \leq 1 - \Lambda_{it}^{[j]}, \quad j = 1, \dots, J+1.$$

The adding-up restriction necessarily induces contemporaneous dependence across equations and would make the covariance matrix singular if all $J+1$ residuals were included. The model therefore uses the J non-base outcomes and does not impose contemporaneous independence across equations.

To state the conditions across time, let $\epsilon_{it} = (\epsilon_{it}^{[1]}, \dots, \epsilon_{it}^{[J+1]})'$. For every $(t, s) \in \mathcal{P}$, we impose

$$\mathbb{E}(\epsilon_{ir}^{[j]} \mid x_{it}, x_{is}, \alpha_i) = 0, \quad r \in \{t, s\}, \quad j = 1, \dots, J+1, \quad (13)$$

and

$$\mathbb{E}(\epsilon_{it} \epsilon_{is}' \mid x_{it}, x_{is}, \alpha_i) = \mathbf{0}. \quad (14)$$

Condition (14) permits arbitrary contemporaneous dependence across outcome equations and requires only conditional cross-time uncorrelatedness for the selected pair. Conditional independence of the complete outcome vectors at t and s , given $(x_{it}, x_{is}, \alpha_i)$, is sufficient but not necessary. As in the univariate case, conditions (13)–(14) need hold only for the time pairs included in $\mathcal{P}$.

Using $\Delta x_{i,t,s} = x_{it} - x_{is}$, define the pairwise transformed residual for non-base outcome j as

$$u_{i,t,s}^{[j]}(\beta^{[j]}) = y_{it}^{[j]} y_{is}^{[J+1]} - \exp(\Delta x_{i,t,s}' \beta^{[j]}) y_{is}^{[j]} y_{it}^{[J+1]}, \quad j = 1, \dots, J. \quad (15)$$

The restrictions above imply the conditional moments

$$\mathbb{E}[u_{i,t,s}^{[j]}(\beta^{[j]}) \mid x_{it}, x_{is}] = 0, \quad j = 1, \dots, J, \quad (16)$$

and the corresponding unconditional moments

$$\mathbb{E}[u_{i,t,s}^{[j]}(\beta^{[j]}) q_{i,t,s}] = 0, \quad j = 1, \dots, J. \quad (17)$$

Here $q_{i,t,s}$ is the $Q \times 1$ instrument vector introduced in the univariate setting; under the initial exogeneity assumptions, $q_{i,t,s} = \Delta x_{i,t,s}$.

Let

$$\beta = \left((\beta^{[1]})', \dots, (\beta^{[J]})' \right)', \quad \mathbf{u}_{i,t,s}(\beta) = \left(u_{i,t,s}^{[1]}(\beta^{[1]}), \dots, u_{i,t,s}^{[J]}(\beta^{[J]}) \right)'$$

The $JQ \times 1$ pairwise system moment contribution is

$$m_{i,t,s}(\beta) = \mathbf{u}_{i,t,s}(\beta) \otimes q_{i,t,s}.$$

Substituting these system contributions into the pooled or stacked individual-level moment functions (11) and (12) yields a GMM estimator analogous to the univariate estimator. The pooled moment function has dimension $JQ \times 1$, while stacking R difference orders gives RJQ moments. The full GMM covariance matrix accommodates dependence among the J outcome equations.

While the semi-elasticity of $y_{it}^{[j]}$ with respect to the k th explanatory variable $x_{k,it}$ depends not only on $\beta_k^{[j]}$ but on all coefficients $\beta_k^{[1]}, \dots, \beta_k^{[J]}$, it remains consistently estimable. Given consistent GMM

estimates $\widehat{\beta}^{[1]}, \dots, \widehat{\beta}^{[J]}$, the average semi-elasticity can be estimated as

$$\widehat{\mathbb{E}}(e_{it,k}^{[j]}) = \widehat{\beta}_k^{[j]} - \sum_{\ell=1}^J \bar{y}^{[\ell]} \widehat{\beta}_k^{[\ell]}, \quad \bar{y}^{[\ell]} = N^{-1}T^{-1} \sum_i \sum_t y_{it}^{[\ell]},$$

for $j = 1, \dots, J$. For the base share, the corresponding estimator is

$$\widehat{\mathbb{E}}(e_{it,k}^{[J+1]}) = - \sum_{\ell=1}^J \bar{y}^{[\ell]} \widehat{\beta}_k^{[\ell]}.$$

2.3 Numerical Assessment

We now report results of a series of simulation experiments designed to demonstrate the feasibility of the estimators introduced above and their adequate finite-sample performance, and to highlight differences between the proposed estimators and other potential estimators.

Univariate fractional dependent variable. We first test the estimator in a setting with a univariate fractional dependent variable, $y_{it} \in [0, 1]$. The data generating processes (DGPs) comprise normally distributed fixed effects α_i and a single regressor x_{it} :

$$\begin{aligned} \alpha_i &\sim N(0, \sqrt{1/2}), \\ x_{it} &= \alpha_i + \sqrt{1/2 - \gamma^2} \xi_{it} + \gamma v_{i,t-1}, \text{ where } \xi_{it} \sim N(0, 1), \quad i = 1, \dots, N, \quad t = 1, \dots, T. \end{aligned}$$

For now, let $\gamma = 0$, which implies that x_{it} is strictly exogenous. The implied correlation between fixed effects and the regressor is independent of T and amounts to about 70 percent ($\rho = 0.5/\sqrt{0.5} \approx 0.707$). The dependent variable is generated as the share of successes of a binomial random variable with success probability Λ_{it} ,

$$y_{it} = \frac{1}{c} \sum_{r=1}^c d_{irt}, \quad d_{irt} = \mathbf{1}\{U_{irt} \leq \Lambda_{it}\}, \quad U_{irt} \stackrel{\text{iid}}{\sim} U(0, 1).$$

Thus, the success probability is the mean of the fractional variable y_{it} . We specify

$$\Lambda_{it} = \Lambda(\beta_0 + \beta_1 x_{it} + \alpha_i),$$

with $\beta_1 = 1$ throughout all DGPs. In the baseline DGP, $N = 1,000$ and $T = 2$, to focus on the incidental parameter problem. The parameter c determines the support points of the dependent variable. For $c = 1$, the DGP is a binary logit. As c increases, y_{it} has support points inside $[0, 1]$. In the baseline case, $c = 10$, which implies $y_{it} \in \{0, 0.1, 0.2, \dots, 1\}$. The parameter β_0 controls the share of bound or limit observations. We specify $\beta_0 = -2, 0, 2$. The first and last cases give substantial shares of observations with $y_{it} = 0$ and $y_{it} = 1$, respectively, while $\beta_0 = 0$ results in a moderate

percentage of both 0 and 1 observations. Figure 1 shows how the distributions for the dependent variable vary with c and β_0 .

We also modify the baseline DGP to consider the case of a violation of strict exogeneity. We let $\gamma = 0.5$, which makes x_{it} a predetermined regressor while preserving its marginal distribution and correlation with the fixed effect. We define the random variable $v_{i,t-1}$, which is now part of x_{it} , as

$$v_{it} = -\frac{1}{\sqrt{c}} \sum_{r=1}^c \Phi^{-1}(U_{irt}).$$

By construction, $v_{it} \sim N(0, 1)$ and is correlated with the contemporaneous CEF error $\epsilon_{it} = y_{it} - \Lambda_{it}$. Consequently, x_{it} is correlated with $\epsilon_{i,t-1}$ but not with current or future CEF errors, and is therefore predetermined. As discussed in Section 2.1, moments based on adjacent periods are invalid in this setting, while moments comparing periods t and $t - 2$ remain valid.

The results for six different settings of the DGP are presented in Table 1. The novel quasi-differenced logit estimator is denoted as “QD” in the table. In all six settings, the estimates are virtually unbiased. With two time periods, the only possible version of the estimator is one based on the first (quasi-) difference, but with $T = 8$ (Panel II), we contrast two versions of the estimator, based on only the first difference (“QD first diff”) and based on all available differences (“QD all diff”). Using all differences increases the efficiency and reduces the estimated root mean squared error (RMSE) by about 15 to 20 percent, depending on the specification. In Panel VI, QD instead uses only the valid second-order pairs $(t, t - 2)$. Its mean estimates range from 0.997 to 1.000 and its RMSE ranges from 0.020 to 0.025, confirming that the estimator remains virtually unbiased when the regressor is predetermined.

To provide a benchmark for the performance of the QD estimator, the simulation shows results from the pooled logit (“Pool”), which ignores the fixed effects. This is the Papke-Wooldridge logit PL cross-section estimator applied to panel data without any modifications. In Panel IV, the DGP does not include fixed effects ($\alpha_i = 0$ for all i), and Pool is consistent. The efficiency loss of using QD, an estimator which can handle potential fixed effects, relative to Pool, which cannot, is substantial, with Pool’s RMSE being only about 40 percent of QD’s. However, in all settings with fixed effects, Pool’s performance is dismal, displaying large biases and an RMSE which is often an order of magnitude larger than QD’s.

Another estimator that we include in the simulation is the Poisson fixed effects pseudo-likelihood estimator (Hausman, Hall and Griliches, 1984) applied to the transformed dependent variable $\tilde{y}_{it} = y_{it}/(1 - y_{it})$. The estimator is denoted as “Pois” in the table. The Pois estimator assumes the conditional expectation function

$$E(\tilde{y}_{it}|x_{it}, \alpha_i) = E\left(\frac{y_{it}}{1 - y_{it}} \middle| x_{it}, \alpha_i\right) = \exp(\beta_0 + \beta_1 x_{it} + \alpha_i),$$

which could therefore be estimated by any fixed effects exponential mean model in principle. Since, in general,

$$E\left(\frac{y_{it}}{1 - y_{it}} \middle| x_{it}, \alpha_i\right) \neq \frac{E(y_{it} \mid x_{it}, \alpha_i)}{1 - E(y_{it} \mid x_{it}, \alpha_i)},$$

Pois is a misspecified benchmark for the binomial DGP considered here. This estimator is related to the [Ramalho, Ramalho and Coelho \(2018\)](#) family of estimators. The transformation can be formed for $y_{it} \in [0, 1)$ or, if y_{it} is appropriately redefined, for $y_{it} \in (0, 1]$, but never for the entire interval $[0, 1]$. Taking the logarithm of $\tilde{y}_{it}$ would make it possible to estimate the model by OLS but would further restrict the possible support of y_{it} to $(0, 1)$. The results in the table show that Pois is extremely sensitive to the share of limit observations, its performance deteriorating drastically when there are more observations with $y_{it} = 1$ in the data. Even when the share of $y_{it} = 1$ is almost negligible (Panel III, $\beta_0 = -2$), this estimator has an RMSE that is much larger than QD's.

Another shortcoming of Pois is that, because it cannot handle both bounds, this estimator cannot be applied to the binary logit case (Panel V): after losing observations at one bound, no variation is left in the dependent variable. For this case, therefore, we consider a fourth estimator, [Chamberlain's \(1980\)](#) conditional logit estimator (Clogit). While this estimator is consistent for the DGP with binary responses $c = 1$ (Panel V), it cannot be implemented in DGPs with fractional responses (Panels I–IV and VI). Taking this estimator as the efficient benchmark for binary responses, the QD suffers an efficiency loss of between 25 and 45 percent in terms of RMSE.

Finally, we consider naïve, brute-force estimation of the pooled model with fixed effects by including a complete set of individual dummy variables ([Greene, 2004](#)), an inconsistent approach that suffers from the incidental parameters problem. We denote this estimator as “PoolFE” in the table. The estimator shows some bias and a larger RMSE than QD in Panels I, III and IV. The incidental parameter bias decreases in T , which can be seen in the improved performance of PoolFE in Panel II which has $T = 8$, and in which PoolFE has a smaller RMSE than QD (first diff), although not than QD (all diff). This favorable performance with $T = 8$ does not carry over when the regressor is predetermined. Holding T and the marginal distribution of x_{it} , including its correlation with α_i , fixed relative to Panel II, the PoolFE mean estimates in Panel VI range from 0.937 to 0.945 and the RMSE ranges from 0.057 to 0.065. By comparison, QD based on the valid second-order pairs has mean estimates between 0.997 and 1.000 and RMSE between 0.020 and 0.025. Thus, PoolFE is sensitive to feedback from past outcome innovations even when its incidental parameter bias is small. The brute-force approach also has two important practical drawbacks. First, its bias in a binary response setting (Panel V) can be very large, as is known ([Todd, 2007](#)) and can be observed in the table. Second, the estimator is computationally intensive, requiring a computing time that, in our implementation, was sixty times as long as that of all the other estimators combined. The limitations of this approach become even more severe when a DGP with a multivariate fractional outcome is considered, as we do next.

Multivariate fractional dependent variable. In our simulation DGP for a multivariate fractional response variable, we specify three outcomes $\mathbf{y}_{it} = (y_{it}^{[1]}, y_{it}^{[2]}, y_{it}^{[3]})'$, with $y_{it}^{[1]} + y_{it}^{[2]} + y_{it}^{[3]} = 1$ and $y_{it}^{[j]} \in [0, 1]$ for $j = 1, 2, 3$. We specify

$$\Lambda_{it}^{[1]} = \frac{\exp(z_{it}^{[1]})}{1 + \exp(z_{it}^{[1]}) + \exp(z_{it}^{[2]})}, \quad \Lambda_{it}^{[2]} = \frac{\exp(z_{it}^{[2]})}{1 + \exp(z_{it}^{[1]}) + \exp(z_{it}^{[2]})},$$

and $\Lambda_{it}^{[3]} = 1 - \Lambda_{it}^{[1]} - \Lambda_{it}^{[2]}$, where

$$z_{it}^{[1]} = \alpha_i^{[1]} + \beta_0^{[1]} + \beta_1^{[1]}x_{it}, \quad z_{it}^{[2]} = \alpha_i^{[2]} + \beta_0^{[2]} + \beta_1^{[2]}x_{it}.$$

For each of c trials, let

$$d_{irt}^{[1]} = \mathbf{1}\{U_{irt} \leq \Lambda_{it}^{[1]}\}, \quad d_{irt}^{[2]} = \mathbf{1}\{\Lambda_{it}^{[1]} < U_{irt} \leq \Lambda_{it}^{[1]} + \Lambda_{it}^{[2]}\}, \quad d_{irt}^{[3]} = \mathbf{1}\{U_{irt} > \Lambda_{it}^{[1]} + \Lambda_{it}^{[2]}\}, \quad U_{irt} \sim U(0, 1)$$

and define

$$y_{it}^{[j]} = \frac{1}{c} \sum_{r=1}^c d_{irt}^{[j]}, \quad j = 1, 2, 3.$$

Thus, the three simulated outcomes lie on the unit simplex in every observation.

The fixed effects in the baseline DGP are generated as

$$\alpha_i^{[1]} = \alpha_i, \quad \alpha_i^{[2]} = -1.5\alpha_i,$$

where α_i is taken from the univariate DGP. The regressor x_{it} is also generated in the same way as before. The coefficients are set as $\beta_1^{[1]} = 1$ and $\beta_1^{[2]} = -1$. The constants determine the share of bound observations. We set $\beta_0^{[1]} = \beta_0^{[2]}$ throughout, and consider the values of -2, -1 and 0.

The results for the six settings are shown in Table 2. Here, the benchmark Pool estimator is a system GMM estimator based on the two conditional-mean residuals $y_{it}^{[1]} - \Lambda_{it}^{[1]}$ and $y_{it}^{[2]} - \Lambda_{it}^{[2]}$. It is consistent when the outcome fixed effects are absent, as confirmed by Panel IV, where its mean estimates are close to the true values and its RMSE is smaller than QD's. In all designs with outcome fixed effects, however, Pool displays substantial omitted-fixed-effect bias. In contrast, the QD estimator performs well in all the designs. Panels I–V use the strictly exogenous version of the regressor. Across Panels I–IV, the absolute bias of either slope estimate does not exceed 0.017 and the RMSE ranges from 0.019 to 0.077. Panel II again shows the efficiency gain from exploiting additional time-pair moments: using all available pairwise quasi-differences instead of only adjacent-period pairs reduces the RMSE by approximately 10 to 20 percent. Panel V, where $c = 1$ and therefore the model is a fixed effects multinomial logit, is considerably more demanding. QD remains much less biased than Pool, but displays some finite-sample bias and greater sampling variation than in the fractional-response designs. Panel VI provides the predetermined-regressor experiment for the multivariate case. When QD uses

the valid second-order pairs, the mean estimates range from 0.999 to 1.001 for $\beta_1^{[1]} = 1$ and from -1.006 to -0.995 for $\beta_1^{[2]} = -1$. The corresponding RMSEs range from 0.024 to 0.030. Its performance is therefore comparable to that obtained from adjacent-period pairs under strict exogeneity in Panel II. This confirms that the multivariate QD estimator remains virtually unbiased under predetermination, provided that the time pairs used to construct the moments are chosen appropriately.

In this multivariate setting, none of the other estimators used in the univariate simulations is straightforward to implement. The only estimator that could be used in some settings is the brute-force fixed effects estimator, PoolFE. With three outcomes and $N = 1,000$ panel units, however, this approach requires two individual fixed effects for each panel unit, or $JN = 2,000$ fixed-effect parameters, in addition to the slope and intercept parameters. This exceeds the number of observations when $T = 2$ and makes the approach infeasible. Although PoolFE could in principle be implemented when $T = 8$, we exclude it because of its computational requirements.

Overall, the simulations show that, across a broad range of univariate and multivariate fractional-response designs, the QD estimator has little finite-sample bias and relatively low RMSE. They also demonstrate that the estimator can accommodate predetermined regressors by selecting time-pair moments that respect the weaker exogeneity condition.

3 The effect of NDIS support on households' time allocation

In this section, we apply our QD estimator to study the effect of the NDIS on the time allocation of households with at least one eligible member. The QD estimator is particularly suited to this application because time use in a given week is constrained to 168 hours, and accordingly, the allocation of time across different activities can be naturally expressed in fractional units.

3.1 A new disability support scheme

The NDIS is the first national approach to provide individualized funding for disability supports. It began in 2013 with a few trial sites, with broader roll-out occurring from 2016 until 2020 when full national coverage was reached. Figure 2 shows the progressive geographic roll-out over time. The introduction of the NDIS represented a major change to Australia's social support system for people with disabilities. The new system shifted from a state-based funding and provision model to a market-based fee-for-service approach, which allows people with a disability to have direct control over the services they need. The role of the NDIS is that of assessor and funder. The average payment to a recipient in 2024-2025 was AUD 64,400 – higher than the median Australian personal income ([National Disability Insurance Agency, 2024](#); [Australian Bureau of Statistics, 2025](#)). While the Australian NDIS is part of a trend towards individualized funding for self-directed care in Western

Europe, North America, and other OECD countries ([Needham, 2013](#)), two features make the NDIS unique. Where other countries have introduced more limited market-based disability services over longer timescales, the NDIS stands out for the broadness of its scope—the NDIS is committed to funding most non-medical support that is “reasonable and necessary” and not provided by other mainstream services—and for the fast rollout of the scheme which replaced the old support system within less than a decade ([Olney and Dickinson, 2019](#)).

The NDIS aims to provide reasonable and necessary non-medical supports to eligible Australians who enter the scheme under the age of 65 years, with a permanent and significant disability (intellectual, cognitive, neurological, sensory, physical or psychosocial). An explicit objective of the NDIS is to support the independence and social and economic participation of people with disability ([Commonwealth of Australia, 2013](#)). Despite this objective, the true impact of the NDIS on the allocation of household time remains unclear. The NDIS may lead to increased employment opportunities for recipients, for example, through better support to travel to work, a greater emphasis on workplace skills and positive behavior strategies, and funding for personal care supports at work. On the other hand, if recipients were working for the primary purpose of funding their own support services, the NDIS support package may allow them to reduce their employment hours. For other household members of NDIS recipients, NDIS-funded services and supports may partially substitute for their informal care, which can free time that can be allocated towards paid work (or leisure activities). However, it is also possible that the market-based approach places a greater burden on household members to organize and co-ordinate the different support services, which could ultimately lead to an increase in caregiving time. This *ex-ante* ambiguity in how the NDIS may impact the allocation of time among its recipients and their family members warrants a careful empirical examination.

3.2 Household Income and Labour Dynamics in Australia Survey (HILDA)

To estimate the effect of NDIS on time use, we use information from HILDA, a large yearly survey representative of Australian households, which started in 2001 and is ongoing. It is funded by the Department of Social Services and managed by the Melbourne Institute at the University of Melbourne. For more information on the sampling design of HILDA, see [Watson and Wooden \(2002\)](#). HILDA’s cross-section encompasses about 17,000 individuals. We use information from waves 2002–2023.

From 2017, HILDA includes a binary question on NDIS support, asking of every individual in the household (including children) “Has [name] received an agreed package of support through the NDIS at any time during the last 12 months?” This is our key explanatory variable.

Our sample consists of all respondents aged 15 and over who live in a household that at some point in our observation period receives NDIS support. The sample contains 14,953 observations from 1,305 individuals. Breaking down these individuals into NDIS-recipient or other household member roles,

we find about 29.3 percent of individuals are an NDIS recipient in the household (“own”) and about 76.1 percent are other members of the household who are non-NDIS participants (“other”). These add up to more than 100% because about 5.4 percent both receive NDIS support and have a household member who is an NDIS recipient.

In every wave, HILDA asks the question “Does anyone here have any long-term health condition, disability or impairment that has lasted, or is likely to last, 6 months or more; restricts everyday activity; and cannot be corrected by medication or medical aids?” We use the response to this question to broadly indicate disability status. We expect that a proportion of individuals who report having a disability will become an NDIS-recipient following the roll-out of the NDIS in their local area, but not all people with a disability will qualify for the NDIS (due to not meeting severity or condition criteria).

Time use is measured in HILDA by asking each member of the household aged 15 years and older to indicate “How much time would you spend on each of the following activities in a typical week”. Respondents then indicate how much time in hours and/or minutes they spend on each of nine categories listed in Table 3. Respondents are provided with the information that their answers cannot exceed 168 hours (the number of hours in one week) and typically will not be greater than 120. Table 3 shows the distribution of this variable and selected control variables for several subsamples in periods before and after a household receives NDIS support.

To operationalize the time use variable and bring it closer to economic models of time allocation, we group the nine categories into three economically meaningful categories in the context of the impact of disability support on time use: time spent on **paid employment** (includes commuting time), time spent on **housework** (includes running errands, housework, outdoor tasks and caring for own and other children), and time spent on **disability care**. In a classical setting, the first category captures the labor supplied by the individual in exchange for income to produce goods or services demanded by the market. The second two, housework and disability care, represent categories where the individual can either provide their own labor to produce these services or use their income to buy these services externally from the market. Finally, the remaining time is often referred to as leisure in economic settings and is thought of as consumption activities or utility-providing experiences which by their nature cannot be bought on the market (such as one’s sleeping time). In the context of people with disability, this remaining category might often include therapeutic activities such as physiotherapy exercises or other forms of self-care, and we refer to it henceforth as **rest**. The distribution of the time use variable is presented in Figure 3.

Figure 3 shows that individuals in households that at some point receive NDIS support on average spend about 18 percent of weekly hours in paid employment when nobody in the household suffers from a disability. The presence of a disability in a household member (without NDIS support) reduces average employment time to 16 percent for other household members without a disability and to 9

percent for members with a disability. When these households have access to NDIS support, the member with a disability works 5 percent, on average.

3.3 Empirical model of the effect of NDIS on time use

We specify variations of the following empirical baseline model

$$\mathbb{E}[y_{it}^{[j]} | \mathbf{x}_{it}, \alpha_i^{[1]}, \dots, \alpha_i^{[J]}] = \frac{\exp(\alpha_i^{[j]} + \mathbf{x}_{it}' \beta^{[j]})}{1 + \sum_{l=1}^J \exp(\alpha_i^{[l]} + \mathbf{x}_{it}' \beta^{[l]})},$$

where $j = \{\text{work, housework, disability care}\}$, and $\mathbf{x}_{it}' = (\text{NDIS-own}_{it}, \text{NDIS-other}_{it}, x_{it}')'$. Thus, the multivariate dependent time use variables $y_{it}^{[j]}$ are the fractions of weekly time spent in paid employment (“work”), housework, and caregiving, as well as the base category fraction of time spent on any other activity, which we refer to as “rest” and which is equal to

$$y_{it}^{[\text{rest}]} = 1 - y_{it}^{[\text{work}]} - y_{it}^{[\text{housework}]} - y_{it}^{[\text{disability care}]}.$$

The key variables of interest are the indicator variable “NDIS-own”, which is equal to one if respondent i received NDIS support in year t , and the indicator variable “NDIS-other”, which is equal to one if in year t someone in respondent i ’s household (other than i) received NDIS support.

The baseline estimation sample comprises individuals 65 years old or younger from waves 2011-2023 in HILDA who live in households that will at some point receive NDIS access. We exclude information from Tasmania (Australia’s smallest state) in the baseline sample as Tasmania did not have a geographically staggered NDIS rollout and therefore it does not offer cross-sectional within-state variation in NDIS eligibility. Covariates x_{it} include year dummies and age². To capture the household structure we further include in x_{it} the number of household members, children aged 0–4, children aged 5–9, children aged 10–14, and dependents aged 15–24. The estimation sample thus covers a time period starting a few years before the rollout of the NDIS, the complete time span of the rollout, as well as a couple of years when the rollout had been completed.

We estimate the model using the proposed QD logit estimator as described in the previous section. This estimator always accounts for individual fixed effects (FE) and is thus based on a before-after comparison within individuals to identify the effect of the NDIS policy. Further, we use an instrumental variable approach (IV-FE) to address the key potential endogeneity of the timing of NDIS access. This time-varying endogeneity in receiving NDIS support is not accounted for by the FE approach. We use the exogenous variation stemming from the NDIS eligibility rollout as the basis for constructing our instrumental variable. Data on NDIS eligibility rollout were obtained from [Ma et al. \(2026\)](#). Our use of this eligibility variation to instrument actual receipt of support, addressing endogenous program take-up, is a further contribution to the NDIS literature and builds on the established use of exogenous

variation in treatment eligibility or access to identify the effects of treatment receipt (e.g. Angrist, 1990; Finkelstein et al., 2012).

Using the rollout as an IV assumes that, from the perspective of potential recipient households, the rollout was as good as random, and that households did not relocate to areas with eligibility. The first requirement is plausible, especially considering that the implementation is conditional on all time-invariant factors. The second requirement is empirically testable since we can observe relocation of households: in our data none of the households moved into an eligible region. This fact possibly reflects the high cost of moving for households with disability, which exceeds the costs of waiting for the region of residence to become eligible for NDIS. We define $eligible_{it}$ as an indicator equal to one if the NDIS had been rolled out in individual i 's area by the date of their HILDA interview at time t , and zero otherwise. We construct this indicator by matching the precise interview date to the rollout date specified in the relevant bilateral state agreements that formed the legal basis for the NDIS rollout. Our excluded instruments are the interactions $eligible_{it} \times disability_own_{it}$ and $eligible_{it} \times disability_other_{it}$, corresponding to the endogenous indicators NDIS-own and NDIS-other, respectively. These instruments allow us to identify potentially different effects of NDIS receipt on recipients and other household members, subject to the rank condition.

We focus on two objects of interest: marginal effects and semi-elasticities of NDIS support. For the average semi-elasticity, the QD estimator provides a simple, consistent estimate as discussed above. For marginal effects, we estimate the marginal effect *at the average* y_{it} , which does not rely on the α_i and can therefore be estimated consistently.

3.4 Results

The baseline estimation results are presented in Table 4. Panel I displays the estimated coefficients, and Panels II and III display the derived marginal effects and semi-elasticities. Overall, the results are broadly similar when looked at through the lens of either marginal effects or semi-elasticities. The FE estimates tend to find effects of NDIS across outcomes that are smaller in absolute magnitude than the IV estimates. The IV estimates consistently show a negative effect or semi-elasticity of own NDIS receipt with respect to paid work and disability care. The estimates for housework and the rest (remainder) category are positive but not statistically significant. For other household members (non-recipients in an NDIS-supported household), the estimates are reversed, showing a statistically significant positive effect on disability care and a negative effect on the rest category. The effects on paid work and housework, while also positive, are small and statistically insignificant. In all cases, the magnitudes of the effects are larger for NDIS recipients than for other household members.⁷

⁷Appendix Table 8 reports auxiliary fixed-effects regressions relating the eligibility-based instruments to NDIS receipt and time use. While these regressions are not stages of the nonlinear GMM estimator and do not formally assess its identification strength, they do document the variation underlying the empirical design: The eligibility IVs significantly predict their corresponding NDIS receipt indicators, and the fractional time-use regressions show a negative work

Since women and men often display gendered patterns of work and household activities, we ran separate IV-FE time use models by gender (Table 5). The results show that there are indeed gendered effects in terms of the general patterns of how NDIS support affected time use for recipients and household members. For NDIS recipients, we had seen in the previous results that time spent on both (paid) work and (unpaid) disability care was reduced in response to NDIS access. Table 5 shows that the reduction in work stems primarily from male recipients while the reduction in disability care from female recipients. For NDIS family members, we find that the effects of increasing work and disability care (while decreasing the rest category) are found only in the sample of women, while there are virtually no effects in the sample of men.

We next focus on work and disability care, two outcomes which contain a large share of zero values, to explore the effect of NDIS support on the extensive margin. The previous estimates of NDIS receipt on work and disability care were positive for household members. By considering the extensive margin separately, it can be estimated, for instance, if and to what extent these positive effects include non-working household members starting to work in response to NDIS receipt. The two outcomes $P(y_{it}^{[\text{work}]} > 0)$ and $P(y_{it}^{[\text{disability care}]} > 0)$ represent two separate binary panel data models with fixed effects, which one could estimate with Chamberlain’s conditional logit estimator. However, we proceed with the univariate version of our QD estimator, which has the advantage that it can easily accommodate an IV approach.

The results for the IV-FE QD estimates are reported in Table 6 for the entire sample, as well as for women and men separately. The results show that there are some quantitatively important extensive margin effects at play, although in some other cases the estimation is imprecise. A pattern that is clear is that of the effect of NDIS recipients on work. The marginal effect is large and similar for men and women (between 40 and 50 percentage points), suggesting that a substantial part of the effect estimated previously is due to the extensive margin: NDIS recipients use the support to exit the work force. This is consistent with the previous literature that found NDIS participants regarded work as not important to them (Chen et al., 2024). In contrast, there is no clear evidence that there is an extensive margin effect of NDIS recipients on disability care. Thus, the reductions in disability care estimated previously mostly seem to happen along the intensive margin, rather than stopping altogether. Among other household members of NDIS recipients, there is a large, positive, statistically significant effect of NDIS support on the extensive margin of unpaid disability care found in the entire sample. Similar effects are also estimated for women and men separately, although in these subsamples they are no longer statistically significant. There is some evidence for a negative effect on paid work among male other household members, suggesting the previous overall null finding for male other household members on work might hide heterogeneous effects at the extensive and intensive margins.

Next we examine if the age of the NDIS recipient in the household influences how household members re-allocate their time in response to NDIS support. We classify household members into two groups,

coefficient for own-disability×eligible and a positive disability care coefficient for other-disability×eligible.

depending on whether the NDIS recipient is younger than 15 years old or not, and estimate separate effects for these two groups (Table 7).⁸ The broad picture for family members from the baseline model was one of increasing informal disability care in response to receiving a NDIS support package. The results in Table 7 show that this effect comes from families where the NDIS recipient is older (i.e., ≥ 15 years). In families with a younger NDIS recipient (under 15 years), the effects have the opposite signs. This may be because NDIS supports for children, such as therapies and behavioral supports, can substitute for parent-delivered exercises and support, and can be provided during school hours, which may free up parental caregiver time. Family caregivers are still assumed to provide a baseline level of informal disability care (the NDIS does not replace this), and for older NDIS participants, there may be an increase in time spent engaging with providers, for example, identifying preferred providers, coordinating and supporting travel to providers, and managing cancellations.

3.5 Robustness checks

We performed a range of sensitivity checks to test the robustness of our empirical results. First, we compared the results obtained by the new QD estimator on the fractional response model with those of OLS and IV on a linear model (cf. Appendix Table 9). The linear model provides a direct estimate of the marginal effect, but not of semi-elasticities. We construct the corresponding semi-elasticity using the (inconsistently) estimated α_i . The results obtained were broadly similar. We also tested the sensitivity of our results to using different sample periods. Both using the entire sample period (from 2002 to 2023, Table 10) and the shorter period of the broad-based NDIS rollout (2016-2023, Table 11) resulted in qualitatively similar estimates. Our baseline QD estimates were obtained using the first four quasi-differences. We checked if using fewer differences changed the results (Table 12), and this was not the case, although the estimates became less precise, as would be expected. In the baseline, we excluded observations from Tasmania from estimation as Tasmania did not have a geographically staggered NDIS rollout. However, including Tasmania in the estimation sample did not change the results (Table 13). We also restricted the sample to observations where the respondent was at least 25 years old, to exclude individuals who might be still in education (Table 14). The results are robust to this check, with the negative effect on work being somewhat larger in magnitude.

As a ‘within’-type estimator, our approach relies on the change in NDIS status within individuals over time to estimate the effect of the NDIS policy variables. Our baseline sample is restricted to individuals whose NDIS status is observed to change in the observation period. While not directly contributing to the identification of the NDIS policy effect, individuals without a change in NDIS status could potentially contribute to the efficiency of the estimation and to the identification of the coefficients on covariates. In a final robustness check, we extended the estimation sample to the full

⁸In the few cases where there is more than one NDIS recipient, we use the age of the youngest recipient. The age-15 cut-off is somewhat arbitrary, but it is an age that can signify a young person’s transition to independence, and it is the median age of NDIS recipients for respondents in NDIS-supported households in our data, giving us two equally sized groups.

HILDA panel, that is, we also included all observations that do not change NDIS status ($n = 90,215$, Table 15). The statistical significance improves for some coefficients, and the magnitude of some estimates increases, such as the negative effect of NDIS support on recipients' working hours. Overall, the results are not affected substantially, demonstrating the robustness of our baseline findings.

4 Conclusions

This paper develops a fixed-effects estimation framework for fractional-outcome panel models with a logit conditional mean. The central result is that pairwise quasi-differencing yields moment conditions from which the individual effects have been eliminated. These moments permit consistent estimation with a fixed number of periods, without specifying a distribution for the individual effects or restricting their association with the covariates. The framework accommodates endogenous explanatory variables through instrumental variables and extends to multivariate outcomes whose shares sum to one. It therefore brings together features that are often required jointly in applied work: bounded outcomes, adding-up restrictions, correlated individual heterogeneity, and time-varying sources of endogeneity.

The approach also offers practical advantages for estimation and interpretation. Observations at zero and one require no adjustment, and binary and multinomial discrete outcomes are included as special cases. Average semi-elasticities can be estimated using the slope coefficients and observed outcome averages, including averages for relevant subgroups, without recovering the individual effects. The Monte Carlo experiments show little bias in the fractional-outcome designs considered and demonstrate that combining additional valid time pairs can improve precision. They also illustrate that efficiency depends on the setting and on whether alternative estimators can exploit additional features of the outcome distribution.

The choice of moment conditions remains central to implementation. The required conditional-mean and cross-period covariance restrictions apply to the time pairs used in estimation. Researchers can therefore select pairs to reflect the temporal dependence and endogeneity present in an application. The predetermined-regressor example illustrates how suitably spaced pairs can remain valid when adjacent pairs do not. For endogenous choice or participation variables, instrumental-variable estimation requires orthogonality conditions for the quasi-differenced residual. These features make the framework adaptable while retaining an explicit connection between the identifying assumptions and the moments used for estimation.

Our application examines how Australia's National Disability Insurance Scheme affects time allocation among recipients and other household members. Using HILDA data and instruments based on the geographic rollout of eligibility, we find that recipients reduce time in paid work and in providing disability care. Other household members increase disability care and reduce time devoted to remaining activities. The care response also differs with the recipient's age: estimated care time among household

members falls for recipients younger than 15 and rises for recipients aged 15 or over. These patterns show that the effects of funded disability support extend beyond recipients' employment outcomes and need not move informal care in the same direction across households. Jointly examining several uses of time, while distinguishing recipients from other household members, provides a fuller account of the adjustments associated with the reform.

References

- Ahn, Seung C. and Peter Schmidt. 1995. "Efficient Estimation of Models for Dynamic Panel Data." *Journal of Econometrics* 68(1):5–27.
- Ahn, Seung C. and Peter Schmidt. 1997. "Efficient Estimation of Dynamic Panel Data Models: Alternative Assumptions and Simplified Estimation." *Journal of Econometrics* 76(1–2):309–321.
- Angrist, Joshua D. 1990. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records." *American Economic Review* 80(3):313–336.
<https://www.jstor.org/stable/2006669>
- Arellano, M and S Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies* 58:277–297.
- Arellano, Manuel and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-Components Models." *Journal of Econometrics* 68(1):29–51.
- Australian Bureau of Statistics. 2025. "Personal Income in Australia, 2022–23." Released 14 November 2025; accessed 22 August 2026.
<https://www.abs.gov.au/statistics/labour/earnings-and-working-conditions/personal-income-australia/latest-release>
- Australian Government. 2025. Budget Strategy and Outlook 2025–26. Budget Paper 1 Commonwealth of Australia.
https://archive.budget.gov.au/2025-26/bp1/download/bp1_2025-26.pdf
- Azoulay, P, D Waverly and T Stuart. 2009. "The impact of academic patenting on the rate, quality and direction of (public) research output." *Journal of Industrial Economics* 57:637–676.
- Bates, Michael D., Leslie E. Papke and Jeffrey M. Wooldridge. 2024. "Non linear correlated random effects models with endogeneity and unbalanced panels." *Econometric Reviews* 43(9):713–732.
- Bedard, K and O Deschênes. 2006. "The long-term impact of military service on health: Evidence from World War II and Korean War veterans." *American Economic Review* 96:176–194.
- Bonhomme, Stéphane. 2012. "Functional Differencing." *Econometrica* 80(4):1337–1385.

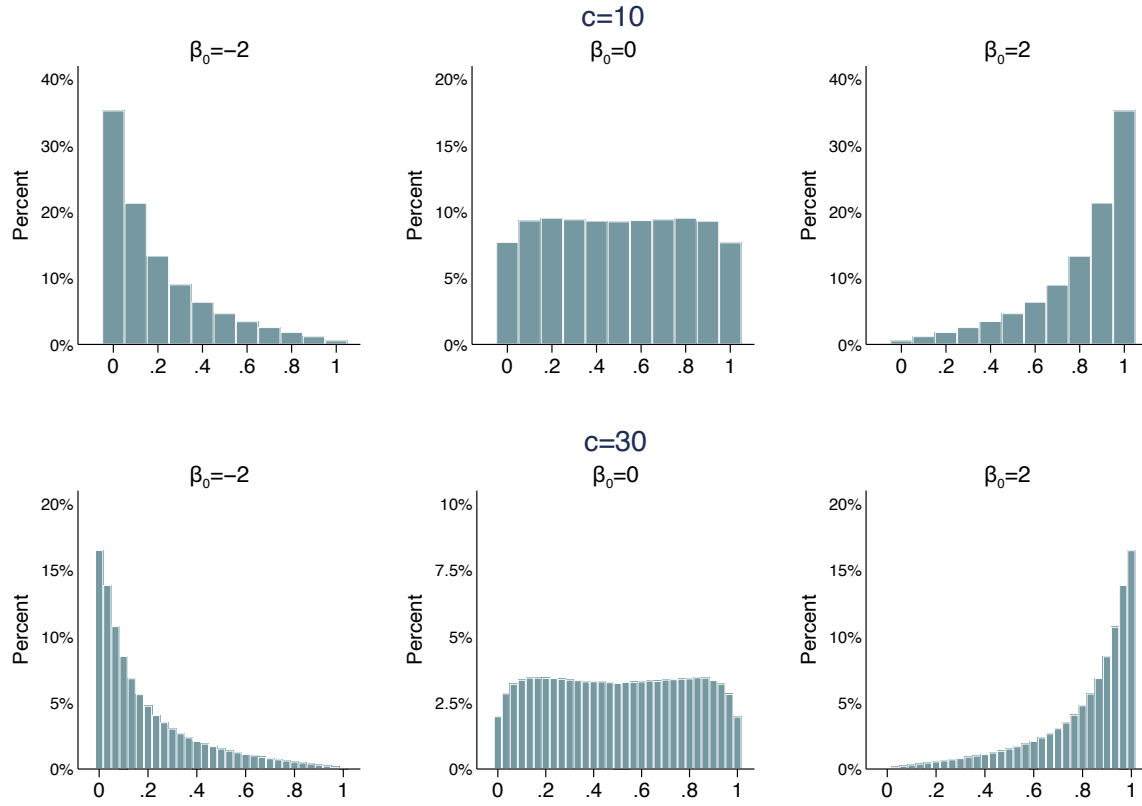
- Chamberlain, G. 1980. "Analysis of covariance with qualitative data." *Review of Economic Studies* 47:225–238.
- Chen, Gang, Dennis Petrie, Gwynnyth Llewellyn, Julie Ratcliffe, Kim Bulkeley, Samia Badji, Maame Esi Woode, Raelene West, Keran Howe, Monique Hines, Jan Abel Olsen, Bernice Hua Ma, Gozde Aydin, Anthony Harris and on behalf of the Disability Wellbeing Index (DWI) Research Team. 2024. *Disability Wellbeing Index: Items Development & Scoring Algorithms*. Melbourne: Monash University.
<https://doi.org/10.26180/30228757.v1>
- Commonwealth of Australia. 2013. "National Disability Insurance Scheme Act 2013 (Cth)." . No. 20, 2013; accessed 22 August 2026.
<https://www.legislation.gov.au/C2013A00020/latest/text>
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker and Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127(3):1057–1106.
- Greene, WH. 2004. "The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects." *Econometrics Journal* 7:98–119.
- Greene, William H. 2012. *Econometric Analysis*. Seventh Edition, International ed. Essex, UK: Pearson Education Limited.
- Hausman, J, BH Hall and Z Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R & D Relationship." *Econometrica* 52:909–938.
- Hausman, JA and GK Leonard. 1997. "Superstars in the National Basketball Association: Economic value and policy." *Journal of Labor Economics* 15:586–624.
- Kemp, GCR and JMC Santos Silva. 2016. "Partial effects in fixed-effects models." *22nd United Kingdom Stata Users' Group, Stata Users Group* .
- Kitazawa, Y. 2012. "Hyperbolic Transformation and Average Elasticity in the Framework of the Fixed Effects Logit Model." *Theoretical Economics Letters* 2:192–199.
- Ma, Bernice Hua, Samia Badji, Dennis Petrie and Gang Chen. 2026. "The Impact of Enhancing Social Care on Healthcare Use for People With Disability: Evidence From Australia." *Health Economics* 35:212–228.
<https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.70055>
- Ma, Bernice Hua, Samia Badji, Gang Chen and Dennis Petrie. 2025. "Exploring the impact of a personalised disability reform on people with disability and their primary carers: Evidence from the

- Australian National Disability Insurance Scheme.” *PLOS ONE* 20(5):e0321377.
<https://doi.org/10.1371/journal.pone.0321377>
- Machado, MP. 2004. “A consistent estimator for the binomial distribution in the presence of incidental parameters: an application to patent data.” *Journal of Econometrics* 119:73–98.
- Mullahy, J. 2015. “Multivariate fractional regression estimation of econometric share models.” *Journal of Econometric Methods* 4:71–100.
- Mullahy, J. 2024. “Analyzing health outcomes measured as bounded counts.” *Journal of Health Economics* 95:102875.
- Mullahy, J and SA Robert. 2010. “No time to lose: Time constraints and physical activity in the production of health.” *Review of Economics of the Household* 8:409–432.
- National Disability Insurance Agency. 2024. “National Quarterly Performance Dashboard: 30 June 2024.”. Accessed 22 August 2026.
<https://www.ndis.gov.au/media/7203/download?attachment=>
- Needham, C. 2013. The Boundaries of Budgets: Why Should Individuals Make Spending Choices about Their Health and Social Care? In *The Three Sector Solution: Delivering Public Policy in Collaboration with Not-for-profits and Business*, ed. J Butcher and D Gilchrist. Australia: ANU Press pp. 319–336.
- Neyman, J and EL Scott. 1948. “Consistent estimates based on partially consistent observations.” *Econometrica* 16:1–32.
- Olney, S and H Dickinson. 2019. “Australia’s New National Disability Insurance Scheme: Implications for Policy and Practice.” *Policy Design and Practice* 2:275–290.
- Papke, LE and JM Wooldridge. 1996. “Econometric methods for fractional response variables with an application to 401(k) plan participation rates.” *Journal of Applied Econometrics* 11:619–632.
- Papke, LE and JM Wooldridge. 2008. “Panel data methods for fractional response variables with an application to test pass rates.” *Journal of Econometrics* 145:121–133.
- Productivity Commission, PC. 2011. “Disability Care and Support: Productivity Commission Inquiry Report.” *Report no. 54* pp. <https://www.pc.gov.au/inquiries/completed/disability-support/report>.
- Ramalho, EA, JJS Ramalho and JMR Murteira. 2011. “Alternative estimating and testing empirical strategies for fractional regression models.” *Journal of Economic Surveys* 25:19–68.
- Ramalho, Esmeralda A., Joaquim J. S. Ramalho and Luís M. S. Coelho. 2018. “Exponential Regression of Fractional-Response Fixed-Effects Models with an Application to Firm Capital Structure.” *Journal of Econometric Methods* 7(1):20150019.

- Santos Silva, JMC, S Tenreyro and K Wei. 2014. “Estimating the extensive margin of trade.” *Journal of International Economics* 93:67–75.
- Tan, Garyn. 2025. “The Labour Market Effects of Insurance for Disability Supports.” Available at SSRN: <https://ssrn.com/abstract=5186362>.
- Todd, P. 2007. “Panel Data: Fixed Effects, Random Effects, Dynamic Panel Data models.” *manuscript, University of Pennsylvania* .
- Wagner, J. 2001. “A note on the firm size–export relationship.” *Small Business Economics* 17:229–237.
- Watson, Nicole and Mark Wooden. 2002. The Household, Income and Labour Dynamics in Australia (HILDA) Survey: Wave 1 Survey Methodology. HILDA Project Technical Paper Series 1/02 Melbourne Institute of Applied Economic and Social Research, University of Melbourne. Revised October 2002.
- Winkelmann, R and L Xu. 2021. “Testing the binomial fixed effects logit model, with an application to female labour supply.” *Empirical Economics* 62:1–30.
- Wooldridge, JM. 2010. *Econometric analysis of cross section and panel data*. MIT press.

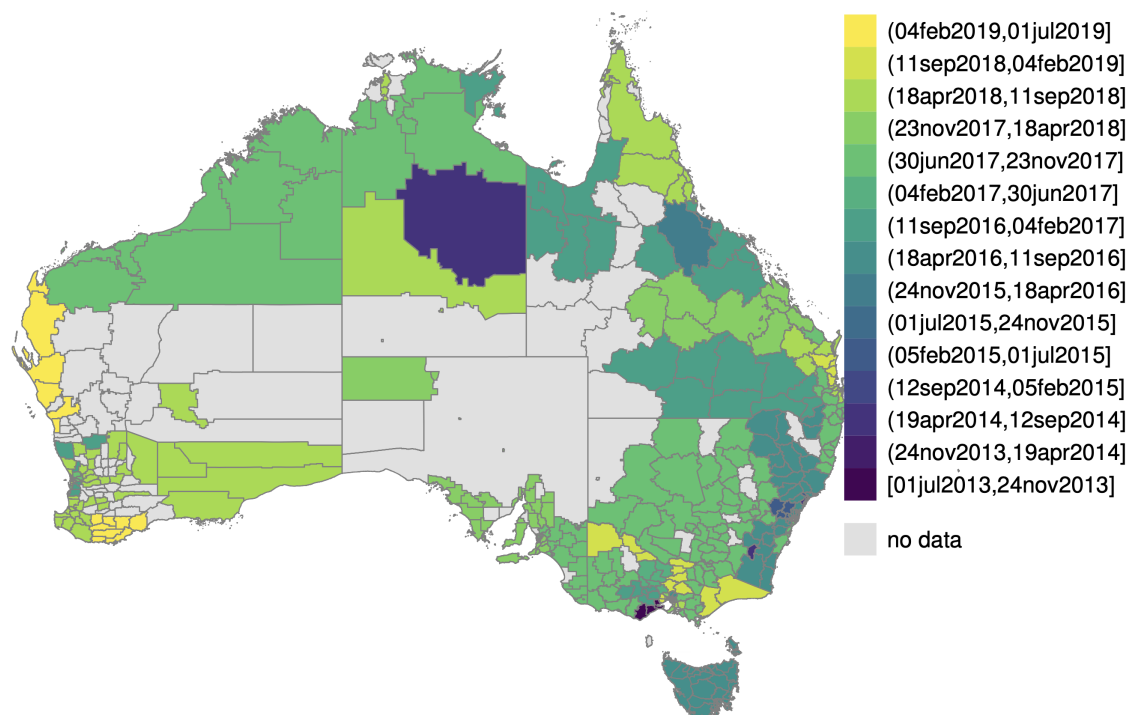
Figures and Tables

Figure 1: MC SIMULATION–UNIVARIATE DGP: DISTRIBUTION OF y_{it} FOR BASELINE DGP AND DGP WITH $c = 30$



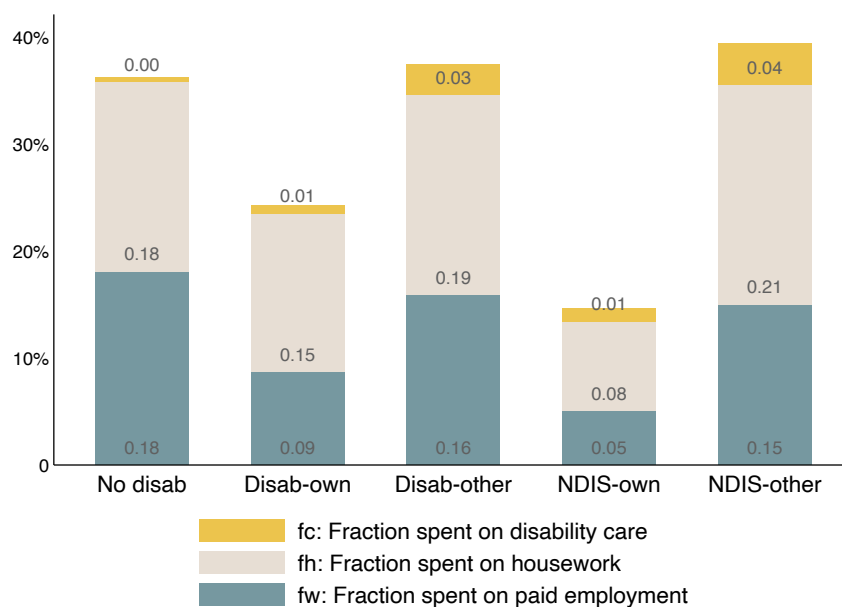
Notes: The top panel shows the probability mass functions of y_{it} when $c = 10$ versus those of y_{it} in the bottom panel when $c = 30$. The figure also compares across 3 different values of $\beta_0 \in \{-2, 0, 2\}$ which control the shape of the distribution and mass of the boundary values.

Figure 2: PROGRESSIVE GEOGRAPHIC ROLL-OUT OF THE NDIS



Source: HILDA 2000-2023, ABS 2024, own calculations.

Figure 3: FRACTION OF WEEKLY TIME ALLOCATED TO PAID EMPLOYMENT, HOUSEWORK AND CAREGIVING



Source: HILDA, 2002-2023, own calculations.

Table 1: MC SIMULATION: UNIVARIATE DGP ($\beta_1 = 1$)

	$\beta_0 = -2$			$\beta_0 = 0$			$\beta_0 = 2$		
	Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
I. <i>Baseline: $N=1,000$, $T=2$, $c=10$, $\rho \neq 0$</i>									
Pool	1.438	0.030	0.439	1.430	0.025	0.431	1.438	0.034	0.439
PoolFE	1.051	0.045	0.068	1.061	0.047	0.077	1.050	0.049	0.071
Pois	0.974	0.072	0.076	0.801	0.056	0.207	0.554	0.042	0.448
QD	1.002	0.049	0.049	1.006	0.056	0.056	0.998	0.063	0.063
II. <i>$T=8$, rest equal to baseline</i>									
Pool	1.441	0.018	0.441	1.429	0.017	0.430	1.440	0.017	0.440
PoolFE	1.016	0.016	0.023	1.013	0.015	0.020	1.013	0.017	0.021
Pois	0.956	0.030	0.053	0.737	0.020	0.264	0.479	0.016	0.521
QD (first diff)	1.001	0.026	0.026	0.998	0.021	0.021	0.999	0.024	0.024
QD (all diff)	1.002	0.020	0.020	0.999	0.018	0.018	0.997	0.021	0.021
III. <i>$c=30$, rest equal to baseline</i>									
Pool	1.440	0.023	0.440	1.429	0.019	0.429	1.438	0.025	0.439
PoolFE	1.017	0.025	0.030	1.019	0.025	0.032	1.017	0.027	0.032
Pois	1.035	0.075	0.083	0.943	0.060	0.083	0.721	0.043	0.283
QD	1.001	0.030	0.030	1.001	0.031	0.031	0.998	0.034	0.034
IV. <i>$\alpha_i=0$, rest equal to baseline</i>									
Pool	1.002	0.020	0.021	0.998	0.018	0.018	1.000	0.024	0.024
PoolFE	1.054	0.044	0.070	1.057	0.039	0.069	1.046	0.048	0.067
Pois	1.051	0.057	0.076	0.926	0.044	0.086	0.484	0.039	0.517
QD	1.004	0.053	0.054	0.997	0.049	0.049	0.994	0.062	0.062
V. <i>$c=1$ (binary response variable), rest equal to baseline</i>									
Pool	1.442	0.086	0.450	1.426	0.065	0.431	1.448	0.075	0.454
PoolFE	2.022	0.340	1.077	2.004	0.293	1.046	2.030	0.301	1.073
Clogit	1.011	0.170	0.170	1.002	0.146	0.147	1.015	0.150	0.151
QD	1.031	0.253	0.254	1.015	0.199	0.199	1.027	0.269	0.270
VI. <i>Predetermined regressor, $T=8$, rest equal to baseline</i>									
Pool	1.442	0.017	0.443	1.431	0.016	0.431	1.441	0.017	0.441
PoolFE	0.937	0.016	0.065	0.945	0.012	0.057	0.940	0.016	0.062
Pois	0.875	0.028	0.128	0.686	0.017	0.314	0.450	0.015	0.550
QD (second diff.)	0.997	0.025	0.025	1.000	0.020	0.020	0.999	0.024	0.024

Notes: Cell entries contain mean (Mean), standard deviation (SD) and root mean squared error (RMSE) for $\beta_1 = 1$ estimated over 100 replications for the estimators Pool (pooled logit MLE), PoolFE (pooled logit MLE with fixed effect dummies), Pois (Poisson PLE) and QD (quasi-differenced GMM). QD (first diff), QD (second diff.) and QD (all diff) denote variants of the estimator using only first pairwise quasi-differences, only pairs $(t, t-2)$, and all available pairwise quasi-differences, respectively. Clogit denotes the conditional (fixed effects) logit estimator. The parameter β_0 is the intercept. In all panels, the number of individuals is $N = 1,000$. T indicates the number of time periods, c the number of binomial trials over which the fractional response variable is calculated, and ρ the correlation between α_i and the regressor x_{it} . Thus, for $\alpha_i = 0$, $\rho=0$. In Panel VI, $x_{it} = \mu_{it} + 0.5v_{i,t-1}$, where $\mu_{it} = \alpha_i + 0.5\xi_{it}$, $\xi_{it} \sim N(0,1)$, and $v_{it} \sim N(0,1)$ is correlated with the contemporaneous CEF residual.

Table 2: MC SIMULATION: MULTINOMIAL DGP ($\beta_1 = 1$ AND $\beta_2 = -1$)

		$\beta_0 = -2$			$\beta_0 = -1$			$\beta_0 = 0$		
		Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
I. <i>Baseline: $N=1,000, T=2, c=10, \rho \neq 0$</i>										
Pool	$\hat{\beta}_1$	1.409	0.033	0.410	1.364	0.033	0.365	1.315	0.036	0.317
	$\hat{\beta}_2$	-1.557	0.036	0.558	-1.500	0.036	0.501	-1.444	0.041	0.446
QD	$\hat{\beta}_1$	1.015	0.071	0.072	1.008	0.059	0.059	1.007	0.066	0.066
	$\hat{\beta}_2$	-1.004	0.077	0.077	-1.002	0.064	0.064	-0.993	0.076	0.076
II. <i>$T=8$, rest equal to baseline</i>										
Pool	$\hat{\beta}_1$	1.402	0.018	0.402	1.358	0.018	0.358	1.311	0.019	0.311
	$\hat{\beta}_2$	-1.565	0.026	0.566	-1.509	0.021	0.509	-1.448	0.023	0.449
QD (first diff)	$\hat{\beta}_1$	1.003	0.026	0.026	1.003	0.025	0.025	1.003	0.027	0.027
	$\hat{\beta}_2$	-0.998	0.027	0.027	-0.999	0.022	0.022	-1.002	0.028	0.028
QD (all diff)	$\hat{\beta}_1$	1.002	0.021	0.021	1.001	0.020	0.020	1.000	0.022	0.022
	$\hat{\beta}_2$	-1.001	0.025	0.025	-1.001	0.019	0.019	-0.998	0.023	0.023
III. <i>$c=30$, rest equal to baseline</i>										
Pool	$\hat{\beta}_1$	1.404	0.025	0.405	1.358	0.028	0.359	1.310	0.031	0.312
	$\hat{\beta}_2$	-1.561	0.036	0.562	-1.508	0.031	0.509	-1.447	0.033	0.448
QD	$\hat{\beta}_1$	1.002	0.036	0.035	1.002	0.039	0.039	1.002	0.043	0.043
	$\hat{\beta}_2$	-0.998	0.038	0.038	-1.002	0.039	0.039	-0.994	0.040	0.041
IV. <i>$\alpha_i=0$, rest equal to baseline</i>										
Pool	$\hat{\beta}_1$	1.002	0.024	0.024	1.001	0.023	0.023	1.002	0.025	0.025
	$\hat{\beta}_2$	-0.999	0.025	0.024	-0.999	0.023	0.023	-0.997	0.027	0.027
QD	$\hat{\beta}_1$	1.017	0.075	0.077	1.010	0.064	0.064	1.010	0.068	0.069
	$\hat{\beta}_2$	-0.997	0.075	0.075	-0.994	0.063	0.063	-0.993	0.067	0.067
V. <i>$c=1$ (one multinomial draw per observation), rest equal to baseline</i>										
Pool	$\hat{\beta}_1$	1.422	0.087	0.431	1.379	0.086	0.389	1.320	0.078	0.330
	$\hat{\beta}_2$	-1.566	0.094	0.573	-1.512	0.078	0.518	-1.442	0.093	0.452
QD	$\hat{\beta}_1$	1.090	0.250	0.265	1.062	0.247	0.253	1.042	0.339	0.340
	$\hat{\beta}_2$	-1.041	0.256	0.258	-1.129	0.320	0.344	-1.105	0.365	0.378
VI. <i>Predetermined regressor, $T=8$, rest equal to baseline</i>										
Pool	$\hat{\beta}_1$	1.404	0.021	0.404	1.360	0.022	0.361	1.312	0.022	0.312
	$\hat{\beta}_2$	-1.562	0.025	0.563	-1.503	0.024	0.503	-1.444	0.025	0.445
QD (second diff)	$\hat{\beta}_1$	0.999	0.027	0.027	1.001	0.024	0.024	1.000	0.025	0.025
	$\hat{\beta}_2$	-1.006	0.028	0.029	-1.003	0.026	0.026	-0.995	0.030	0.030

Notes: Cell entries contain the mean (Mean), standard deviation (SD), and root mean squared error (RMSE) for $\beta_1 = 1$ and $\beta_2 = -1$, estimated over 100 replications for Pool and QD. Pool is the pooled system GMM estimator based on the two multinomial conditional-mean residuals and ignores the fixed effects, while QD is the quasi-differenced GMM estimator. QD (first diff), QD (second diff), and QD (all diff) use only adjacent-period pairs, only pairs $(t, t-2)$, and all available pairwise quasi-differences, respectively. The parameter β_0 is the common intercept for the first two outcome equations. In all panels, the number of individuals is $N = 1,000$. T indicates the number of time periods, and c the number of multinomial trials from which the three fractional outcomes are calculated; hence, $y_{it}^{[1]} + y_{it}^{[2]} + y_{it}^{[3]} = 1$ by construction. The parameter ρ denotes the correlation between the fixed effects and the regressor x_{it} . In Panel IV the outcome fixed effects are zero. In Panel VI, $x_{it} = \alpha_i + \sqrt{1/2 - \gamma^2} \xi_{it} + \gamma v_{i,t-1}$ with $\gamma = 0.5$, making x_{it} predetermined; QD therefore uses only the valid pairs $(t, t-2)$.

Table 3: DESCRIPTIVE STATISTICS

	Before NDIS			After NDIS	
	No disability	Disability–own	Disability–other	NDIS–own	NDIS–other
<i>Time use (hours per week spent in indicated activity)</i>					
Paid work	27.28	13.14	19.53	7.94	21.63
Commute	3.13	1.56	2.28	0.85	2.51
Errands	4.09	4.52	4.43	3.58	4.89
Housework	10.33	11.34	11.97	7.76	12.13
Outdoor tasks	2.80	3.23	3.26	2.11	3.33
Caring for own children	12.52	8.44	11.25	4.49	14.17
Caring for other children	0.70	1.26	1.37	0.84	1.18
Volunteering	0.63	0.80	0.78	0.80	0.72
Disability care	0.69	4.05	5.69	2.65	6.45
<i>Socioeconomic variables</i>					
Age	34.94	41.34	38.34	41.71	40.12
No. of children	1.00	0.79	1.20	0.46	1.52
HH income (\$1000)	87.43	48.65	69.47	48.74	96.62
Own income (\$1000)	49.39	35.93	41.96	37.40	58.32
HH size	3.20	2.91	3.84	2.63	4.13
Observations	4,744	4,298	5,286	817	2,123

Note: Data: HILDA23, waves 2002–2023. Column headings distinguish observations before and after NDIS receipt and whether disability or NDIS receipt applies to the respondent (“own”) or another household member (“other”).

Table 4: THE EFFECT OF NDIS SUPPORT ON TIME USE: FRACTIONAL MODEL QD ESTIMATES

	FE estimates				IV-FE estimates			
	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.165*** (0.064)	0.005 (0.045)	-0.202 (0.236)	–	-0.399** (0.156)	0.080 (0.244)	-1.440** (0.650)	–
NDIS – other	0.049** (0.025)	0.031 (0.031)	0.347** (0.145)	–	0.062 (0.052)	0.036 (0.059)	0.813** (0.367)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.019** (0.007)	0.005 (0.007)	-0.004 (0.005)	0.017* (0.009)	-0.043** (0.019)	0.027 (0.035)	-0.032** (0.015)	0.048 (0.035)
NDIS – other	0.004 (0.003)	0.002 (0.004)	0.008** (0.003)	-0.013** (0.005)	0.004 (0.006)	0.000 (0.009)	0.018** (0.008)	-0.022** (0.011)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.140** (0.056)	0.031 (0.037)	-0.176 (0.230)	0.026* (0.014)	-0.327** (0.140)	0.152 (0.200)	-1.368** (0.632)	0.072 (0.053)
NDIS – other	0.029 (0.022)	0.011 (0.025)	0.327** (0.141)	-0.020** (0.008)	0.029 (0.047)	0.002 (0.048)	0.780** (0.358)	-0.033** (0.016)

Notes: HILDA 2011–2023, own calculations. 6,892 observations. Both specifications use QD logit with individual fixed effects (FE). Controls comprise age squared, the number of household members, the numbers of children aged 0–4, 5–9, and 10–14, the number of dependents aged 15–24, and year indicators. NDIS–own indicates receipt of NDIS support by the respondent; NDIS–other indicates receipt by another household member. IV-FE treats both indicators as endogenous, using the excluded instruments $eligible_{it} \times disability_own_{it}$ and $eligible_{it} \times disability_other_{it}$. Here, $eligible_{it}$ indicates NDIS rollout in the respondent’s area by the interview date; disability–own and disability–other indicate disability of the respondent and another household member, respectively. Rest is the base outcome category. Panel II reports marginal effects evaluated at the sample mean outcome shares. Panel III reports average semi-elasticities. Parentheses report robust standard errors clustered by individual. * $p \leq 0.10$; ** $p \leq 0.05$; *** $p \leq 0.01$.

Table 5: THE EFFECT OF NDIS SUPPORT ON TIME USE: SEPARATE MODELS BY GENDER

	IV-FE estimates: Women				IV-FE estimates: Men			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.350* (0.210)	0.115 (0.299)	-1.695** (0.835)	–	-0.443* (0.233)	0.172 (0.402)	0.851 (1.092)	–
NDIS – other	0.171** (0.076)	-0.003 (0.073)	0.914* (0.481)	–	-0.022 (0.070)	0.110 (0.098)	0.779 (0.612)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.029 (0.019)	0.039 (0.052)	-0.047** (0.023)	0.037 (0.049)	-0.069** (0.034)	0.026 (0.044)	0.015 (0.018)	0.028 (0.046)
NDIS – other	0.013* (0.007)	-0.010 (0.013)	0.025* (0.013)	-0.027* (0.016)	-0.008 (0.010)	0.011 (0.010)	0.013 (0.010)	-0.016 (0.015)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.293 (0.197)	0.172 (0.231)	-1.638** (0.813)	0.057 (0.076)	-0.402** (0.200)	0.212 (0.352)	0.892 (1.075)	0.040 (0.068)
NDIS – other	0.129* (0.070)	-0.045 (0.057)	0.872* (0.466)	-0.042* (0.024)	-0.045 (0.060)	0.087 (0.083)	0.756 (0.600)	-0.023 (0.022)

Notes: HILDA 2011–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. Separate IV-FE models are estimated for women and men using the QD logit estimator. Number of observations: 3,676 (women), 3,214 (men).

Table 6: THE EFFECT OF NDIS SUPPORT ON TIME USE: EXTENSIVE MARGIN EFFECTS FOR WORK AND CARE

	IV-FE estimates Entire sample		IV-FE estimates Women		IV-FE estimates Men	
<i>Dependent variables: Binary variables—any time spent on</i>	work: yes	disab. care: yes	work: yes	disab. care: yes	work: yes	disab. care: yes
<i>I. Marginal effects</i>						
NDIS – own	-0.417*** (0.113)	-0.075 (0.171)	-0.433** (0.216)	-0.048 (0.234)	-0.508*** (0.124)	-0.163 (0.256)
NDIS – other	-0.009 (0.113)	0.095** (0.044)	0.013 (0.160)	0.058 (0.048)	-0.256* (0.134)	0.178 (0.120)
<i>II. Semi-elasticities</i>						
NDIS – own	-0.724*** (0.196)	-0.445 (1.007)	-0.751** (0.374)	-0.284 (1.383)	-0.882*** (0.216)	-0.963 (1.511)
NDIS – other	-0.016 (0.197)	0.559** (0.258)	0.023 (0.278)	0.341 (0.285)	-0.445* (0.232)	1.050 (0.710)
No. of observations	6,934	6,942	3,666	3,647	3,232	3,250

Notes: HILDA 2011–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. Separate univariate QD logit models are estimated for indicators of any time spent on work or disability care. Sample sizes are reported in the table.

Table 7: THE EFFECT OF NDIS SUPPORT ON TIME USE: YOUNGER AND OLDER FAMILY MEMBER NDIS RECIPIENTS

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.166*** (0.064)	0.003 (0.045)	-0.209 (0.236)	–	-0.402*** (0.156)	0.078 (0.247)	-1.445** (0.730)	–
NDIS – other (≥ 15 years)	0.054 (0.035)	0.065* (0.038)	0.445*** (0.161)	–	0.102 (0.094)	0.062 (0.130)	1.068** (0.450)	–
NDIS – other (< 15 years)	-0.009 (0.042)	-0.056 (0.055)	-0.490** (0.212)	–	-0.074 (0.089)	-0.043 (0.124)	-1.011*** (0.371)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.019** (0.007)	0.005 (0.007)	-0.004 (0.005)	0.018* (0.009)	-0.044** (0.019)	0.027 (0.036)	-0.032** (0.016)	0.049 (0.036)
NDIS – other (≥ 15 years)	0.003 (0.004)	0.006 (0.005)	0.010** (0.004)	-0.019*** (0.007)	0.007 (0.011)	0.002 (0.019)	0.024** (0.010)	-0.033* (0.020)
NDIS – other (< 15 years)	0.002 (0.005)	-0.006 (0.008)	-0.011** (0.005)	0.015* (0.009)	-0.004 (0.011)	-0.000 (0.018)	-0.022*** (0.008)	0.027 (0.019)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.139** (0.056)	0.029 (0.037)	-0.183 (0.230)	0.026* (0.014)	-0.329** (0.141)	0.151 (0.202)	-1.372** (0.710)	0.073 (0.054)
NDIS – other (≥ 15 years)	0.025 (0.031)	0.036 (0.031)	0.416*** (0.156)	-0.029*** (0.010)	0.053 (0.085)	0.013 (0.106)	1.019** (0.437)	-0.049** (0.030)
NDIS – other (< 15 years)	0.013 (0.038)	-0.034 (0.044)	-0.468** (0.206)	0.022* (0.013)	-0.033 (0.080)	-0.003 (0.102)	-0.970*** (0.360)	0.041 (0.028)

Notes: HILDA 2011–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. NDIS–other effects are estimated separately according to whether the youngest NDIS recipient in the household is younger than 15 or aged 15 or older. Number of observations: 6,877 (FE), 6,849 (IV-FE).

Appendix

Table 8: ELIGIBILITY, NDIS RECEIPT AND TIME USE: AUXILIARY FIXED-EFFECTS REGRESSIONS

	“Reduced form” – Frac FE			“Reduced form” – OLS FE				“First stage” – OLS FE	
	work	housework	care	work	housework	care	rest	NDIS – own	NDIS – other
disability–own $\times$ eligible	-0.085** (0.035)	0.027 (0.033)	-0.109 (0.157)	-0.006* (0.003)	-0.002 (0.004)	-0.000 (0.004)	0.008 (0.006)	0.233*** (0.016)	-0.054*** (0.015)
disability–other $\times$ eligible	0.039 (0.026)	0.016 (0.029)	0.442*** (0.157)	0.002 (0.003)	0.000 (0.004)	0.013*** (0.003)	-0.015*** (0.005)	-0.047*** (0.011)	0.528*** (0.016)
N	6,855			6,855	6,855	6,855	6,855	6,855	6,855
χ^2_6 (p-val)	16.53 (0.011)								

Notes: HILDA 2011–2023, own calculations. Variable definitions, controls, fixed effects, and inference conventions are as in Table 4. The eligibility-based instruments enter as regressors in these auxiliary regressions. The χ^2_6 statistic tests the joint null hypothesis that the six coefficients on the two eligibility-based instruments are zero across the three fractional time-use equations (work, housework and disability care). The corresponding p-value is reported in parentheses.

Table 9: THE EFFECT OF NDIS SUPPORT ON TIME USE: OLS/IV ESTIMATES FROM A LINEAR MODEL

	FE estimates				IV-FE estimates			
	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>Dependent variables: Fraction of time spent on</i>								
<i>I. Coefficients/Marginal effects</i>								
NDIS – own	-0.020*** (0.004)	-0.002 (0.004)	-0.004 (0.004)	0.027*** (0.006)	-0.061*** (0.014)	-0.007 (0.016)	0.005 (0.013)	0.062*** (0.021)
NDIS – other	0.002 (0.003)	0.001 (0.004)	0.009*** (0.003)	-0.012** (0.005)	-0.002 (0.006)	-0.000 (0.007)	0.026*** (0.005)	-0.023** (0.009)
<i>II. Semi-elasticities</i>								
NDIS – own	-0.150*** (0.030)	-0.010 (0.030)	-0.170 (0.150)	0.040*** (0.010)	-0.460*** (0.100)	-0.040 (0.090)	0.230 (0.550)	0.090*** (0.030)
NDIS – other	0.010 (0.020)	0.000 (0.020)	0.390*** (0.130)	-0.020*** (0.010)	-0.020 (0.050)	-0.000 (0.040)	1.110*** (0.240)	-0.030*** (0.010)

Notes: HILDA 2011–2023, own calculations. Variable definitions, controls, instruments, fixed effects, and inference conventions are as in Table 4. Estimates are from linear models using OLS (FE) and 2SLS (IV-FE). Coefficients are marginal effects; semi-elasticities use fitted outcome levels, including the estimated individual effects. Number of observations: 10,398.

Table 10: THE EFFECT OF NDIS SUPPORT ON TIME USE (APPENDIX SPECIFICATION): QD ESTIMATES USING ALL YEARS (2002-2023)

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.168*** (0.064)	0.013 (0.045)	-0.209 (0.241)	–	-0.413*** (0.156)	0.122 (0.251)	-1.485** (0.602)	–
NDIS – other	0.051** (0.025)	0.033 (0.031)	0.319** (0.143)	–	0.069 (0.051)	0.040 (0.059)	0.791** (0.360)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.020*** (0.008)	0.007 (0.007)	-0.004 (0.005)	0.017* (0.009)	-0.048** (0.019)	0.033 (0.036)	-0.030** (0.012)	0.045 (0.035)
NDIS – other	0.004 (0.003)	0.002 (0.004)	0.006** (0.003)	-0.013** (0.005)	0.005 (0.006)	0.001 (0.008)	0.016** (0.007)	-0.022** (0.010)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.142*** (0.056)	0.039 (0.038)	-0.184 (0.235)	0.025* (0.014)	-0.346** (0.139)	0.189 (0.206)	-1.418** (0.589)	0.067 (0.053)
NDIS – other	0.032 (0.022)	0.013 (0.025)	0.299** (0.139)	-0.020** (0.008)	0.036 (0.045)	0.007 (0.048)	0.758** (0.351)	-0.033** (0.016)

Notes: HILDA 2002–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. The estimation period is extended to 2002–2023. Number of observations: 10,011.

Table 11: THE EFFECT OF NDIS SUPPORT ON TIME USE (APPENDIX SPECIFICATION): QD ESTIMATES USING FEWER YEARS (2016-2023)

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.150** (0.066)	-0.016 (0.048)	-0.184 (0.256)	–	-0.342** (0.155)	0.131 (0.253)	-1.511** (0.613)	–
NDIS – other	0.031 (0.026)	0.030 (0.032)	0.382** (0.153)	–	0.063 (0.053)	0.043 (0.059)	0.935** (0.407)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.016** (0.008)	0.002 (0.007)	-0.004 (0.006)	0.018* (0.010)	-0.037** (0.018)	0.032 (0.036)	-0.036** (0.015)	0.040 (0.035)
NDIS – other	0.002 (0.003)	0.002 (0.004)	0.009** (0.004)	-0.012** (0.005)	0.003 (0.006)	0.001 (0.008)	0.022** (0.010)	-0.026** (0.011)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.124** (0.058)	0.011 (0.039)	-0.158 (0.249)	0.027* (0.014)	-0.283** (0.139)	0.190 (0.209)	-1.451** (0.599)	0.060 (0.052)
NDIS – other	0.012 (0.023)	0.011 (0.026)	0.363** (0.149)	-0.019** (0.008)	0.024 (0.048)	0.004 (0.048)	0.896** (0.396)	-0.038** (0.017)

Notes: HILDA 2016–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. The estimation period is restricted to 2016–2023. Number of observations: 4,002.

Table 12: THE EFFECT OF NDIS SUPPORT ON TIME USE (APPENDIX SPECIFICATION): QD ESTIMATES USING ONLY FIRST AND SECOND DIFFERENCES

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.046 (0.061)	-0.001 (0.045)	-0.072 (0.212)	–	-0.247 (0.184)	0.194 (0.323)	1.199 (1.509)	–
NDIS – other	0.054** (0.026)	0.022 (0.032)	0.376** (0.147)	–	0.106* (0.055)	0.047 (0.068)	0.726* (0.377)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.005 (0.007)	0.001 (0.006)	-0.001 (0.005)	0.005 (0.009)	-0.037* (0.022)	0.029 (0.047)	0.027 (0.034)	-0.020 (0.048)
NDIS – other	0.004 (0.003)	0.000 (0.004)	0.008** (0.003)	-0.013** (0.006)	0.009 (0.007)	0.001 (0.010)	0.016* (0.008)	-0.026** (0.012)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.038 (0.053)	0.007 (0.036)	-0.063 (0.206)	0.008 (0.014)	-0.276* (0.166)	0.165 (0.267)	1.170 (1.479)	-0.029 (0.072)
NDIS – other	0.034 (0.023)	0.003 (0.025)	0.356** (0.142)	-0.020** (0.008)	0.066 (0.049)	0.008 (0.055)	0.687* (0.367)	-0.039** (0.018)

Notes: HILDA 2011–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. Estimation uses only the first two quasi-differences, rather than the first four used in the baseline. Number of observations: 7,580.

Table 13: THE EFFECT OF NDIS SUPPORT ON TIME USE (APPENDIX SPECIFICATION): QD ESTIMATES INCLUDING OBSERVATIONS FROM TASMANIA

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.172*** (0.063)	-0.010 (0.044)	-0.198 (0.235)	–	-0.384** (0.156)	0.085 (0.248)	-1.455** (0.631)	–
NDIS – other	0.050** (0.025)	0.035 (0.031)	0.348** (0.145)	–	0.050 (0.051)	0.018 (0.058)	0.783** (0.366)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.019*** (0.007)	0.003 (0.006)	-0.004 (0.005)	0.019** (0.009)	-0.042** (0.019)	0.027 (0.036)	-0.031** (0.014)	0.046 (0.035)
NDIS – other	0.004 (0.003)	0.003 (0.004)	0.007** (0.003)	-0.014*** (0.005)	0.003 (0.006)	-0.002 (0.008)	0.017** (0.008)	-0.018* (0.011)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.143*** (0.055)	0.019 (0.037)	-0.169 (0.229)	0.029** (0.013)	-0.316** (0.140)	0.153 (0.203)	-1.387** (0.615)	0.068 (0.053)
NDIS – other	0.030 (0.022)	0.014 (0.025)	0.327** (0.141)	-0.021*** (0.008)	0.023 (0.046)	-0.010 (0.048)	0.756** (0.357)	-0.027* (0.016)

Notes: HILDA 2011–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. The estimation sample includes observations from Tasmania. Number of observations: 7,242.

Table 14: THE EFFECT OF NDIS SUPPORT ON TIME USE (APPENDIX SPECIFICATION): USING ONLY OBSERVATIONS WITH AGE ≥ 25

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.204*** (0.071)	0.008 (0.046)	-0.123 (0.246)	–	-0.570*** (0.164)	0.168 (0.288)	-0.939 (1.196)	–
NDIS – other	0.031 (0.025)	0.052 (0.032)	0.423*** (0.147)	–	0.034 (0.051)	0.055 (0.059)	1.197*** (0.358)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.025*** (0.009)	0.008 (0.007)	-0.002 (0.006)	0.019* (0.010)	-0.071*** (0.022)	0.047 (0.045)	-0.023 (0.031)	0.046 (0.046)
NDIS – other	0.001 (0.003)	0.005 (0.005)	0.011*** (0.004)	-0.016*** (0.006)	-0.002 (0.006)	0.002 (0.009)	0.030*** (0.009)	-0.030*** (0.011)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.173*** (0.061)	0.039 (0.038)	-0.093 (0.238)	0.031* (0.016)	-0.497*** (0.156)	0.241 (0.231)	-0.866 (1.159)	0.073 (0.072)
NDIS – other	0.005 (0.022)	0.026 (0.025)	0.397*** (0.142)	-0.026*** (0.009)	-0.013 (0.046)	0.008 (0.047)	1.149*** (0.348)	-0.047*** (0.018)

Notes: HILDA 2011–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. The estimation sample is restricted to respondents aged 25 or older. Number of observations: 5,674.

Table 15: THE EFFECT OF NDIS SUPPORT ON TIME USE (APPENDIX SPECIFICATION): QD ESTIMATES USING THE FULL SAMPLE (INCLUDING NON-NDIS HOUSEHOLDS)

	FE estimates				IV-FE estimates			
<i>Dependent variables: Fraction of time spent on</i>	work	housework	disab. care	rest	work	housework	disab. care	rest
<i>I. Coefficients</i>								
NDIS – own	-0.178*** (0.063)	-0.016 (0.045)	-0.154 (0.239)	–	-1.514*** (0.173)	-0.316 (0.314)	-2.780*** (0.508)	–
NDIS – other	0.035 (0.022)	0.023 (0.029)	0.337*** (0.128)	–	0.381*** (0.129)	0.041 (0.091)	2.096*** (0.566)	–
<i>II. Marginal effects</i>								
NDIS – own	-0.020*** (0.007)	0.002 (0.007)	-0.003 (0.005)	0.020** (0.009)	-0.158*** (0.021)	0.000 (0.045)	-0.055*** (0.011)	0.213*** (0.043)
NDIS – other	0.002 (0.003)	0.001 (0.004)	0.007*** (0.003)	-0.011** (0.005)	0.037** (0.015)	-0.011 (0.013)	0.045*** (0.012)	-0.070*** (0.019)
<i>III. Semi-elasticities</i>								
NDIS – own	-0.149*** (0.055)	0.014 (0.038)	-0.124 (0.233)	0.030** (0.013)	-1.196*** (0.156)	0.002 (0.256)	-2.461*** (0.495)	0.318*** (0.064)
NDIS – other	0.019 (0.019)	0.007 (0.023)	0.321*** (0.125)	-0.016** (0.007)	0.276** (0.112)	-0.063 (0.076)	1.991*** (0.552)	-0.105*** (0.028)

Notes: HILDA 2002–2023, own calculations. Unless otherwise stated, variable definitions, controls, effect measures, and inference conventions are as in Table 4. The estimation sample includes non-NDIS households. Number of observations: 90,215.