Interactive Fixed Effects and Compositional Changes with Application to Immigration Policy and Remittances

Matloob Piracha * Arkadiusz Szydłowski † Massimiliano Tani ‡

August 17, 2026

Abstract

We compare remittance behaviour of two cohorts who entered Australia before and after the 1996 immigration policy change. To address the challenge of evaluating the impact of policy with two distinct migrant samples, we employ an interactive fixed effects model for untreated potential outcomes. We propose a method to deal with compositional changes in this model under the assumption of no change in the mean of fixed effects and derive the asymptotic variance of the estimator. We find that the policy did not have significant effect on remittance behaviour.

Keywords: difference in differences, interactive fixed effects, composition change, immigration, remittances.

JEL Classification: C13, C23, F22, F24, J61.

Data availability statement: Data is available from the Australian Bureau of Statistics.

Funding statement: No funding was received for this paper.

Conflict of interest disclosure: There is no conflict of interest.

Word count: 8961

---

* University of Kent, GLO and IZA. Corresponding author: M.E.Piracha@kent.ac.uk

† University of Kent, A.Szydlowski@kent.ac.uk

‡ University of New South Wales and IZA, m.tanibertuol@unsw.edu.au

1 Introduction

The first objective of the paper is to introduce an econometric framework that allows policy analysis under both interactive fixed effects and compositional changes in covariates in both treatment and control group. This is motivated by our empirical problem – evaluating the effect of 1996 immigration policy reform in Australia using the Longitudinal Survey of Immigrants to Australia (LSIA) data.

The LSIA data is divided into two cohorts. The first cohort entered Australia between 1993 and 1995, just before the policy change (LSIA1), while the second one entered in 1999-2000, after the policy change (LSIA2).[1] In each cohort we have two groups of migrants, the non-family and non-humanitarian group (the treatment group) and the family, employer nominated and humanitarian migrants (the control group).

Since there are two distinct cohorts separated by 4-7 years, they face different job market realities and macroeconomic conditions, which calls for allowing both observed and unobserved characteristics to have different effects between these cohorts, rather than using a simple difference-in-difference approach. Thus, we employ an interactive fixed effect model for untreated outcomes as in Gobillon and Magnac (2016) and Callaway and Karami (2023) (CK23, henceforth).

Migrants in Cohort 1 have different observable characteristics than those who entered the country after the policy change (Cohort 2), since the policy change was partly instituted to select more educated migrants with better English knowledge. Hence, as our main methodological contribution we extend CK23 model to deal with compositional changes in the covariates. To do that, we assume that we only have a change in a composition of the observed covariates, whereas the mean of the unobserved fixed effects is constant across time, and show that the CK23 estimator applies after reweighting the moment conditions by conditional probabilities of being included in the sample at time t.

[1] In other words, each cohort has a short panel data, with three periods for Cohort 1 and two periods for Cohort 2.

As a result, we obtain a three step estimator of the average treatment effect on the treated (ATT), where the first step is the (multinomial) logit estimation of probability of being included in the sample at time t. We derive the asymptotic variance of that estimator.

Beyond the standard difference-in-differences model the literature offers several alternatives to estimate the ATT with repeated cross-sections (or repeated panels). In their seminal paper, Heckman et al. (1997) suggest that the conditional difference-in-differences (DDM) is an effective method to evaluate the effects of a programme under the assumption of conditional parallel trend (see also Abadie (2005)).

In order to deal with compositional changes in that model Hong (2013) performs DDM based on a two-dimensional propensity score. Sant'Anna and Xu (2026) extend this to a fully nonparametric conditional D-in-D setup. Since we do not use the conditional parallel trends assumption but assume interactive fixed effects, our model is not nested within these approaches.

Our methodology contributes to the work in which interactive fixed-effects are used to account for the influence of unobserved time-varying factors, for instance, Bai (2009), Abadie et al. (2012), Gobillon and Magnac (2016), Ouyang and Peng (2015) and Abadie et al. (2015). However, none of these papers deals with compositional changes.

The second objective of this paper is to contribute to the literature on the remittance behaviour of migrants. More precisely, using longitudinal data, we analyse the effect of a change in immigration policy on the probability (extensive margin) and amount (intensive margin) of remittances.

Several aspects could influence the remittance flows from migrants to their families and friends in the sending countries. These range from repaying of loans to fund migration costs, altruism towards those who remain in the country of origin or indeed because of selfish reasons to curry favour with those remaining back home in case migration turns out to be a failure (i.e., equivalent to taking insurance against bad economic outcome). Although migrant remittances have received widespread attention because of their importance for the economies of developing nations, few studies have examined the conditions that affect migrants remittance pattern. These could be determined by the underlying motivations to remit. For instance, in the presence of uncertainty about their legal status and/or labour market conditions, migrants could potentially remit more because of the insurance motive rather than, say, the altruistic motive (see Amuedo-Dorantes and Pozo (2006); Piracha and Zhu (2012); Batista and Umblijs (2016)). These motivations could indeed be affected by the change in conditions in the host country, particularly a change in the immigration policy as that has an implication for immigrants' legal status.

Furthermore, immigration policy could affect the characteristics of immigrants who enter the country and affect their remittance behaviour. For instance, most of the migrant-receiving countries have instituted stringent immigration policies to attract relatively higher skilled migrants, which has raised concerns among many that the resultant brain drain will not be properly compensated for, as the more educated tend to remit less. There are a number of theoretical arguments for this: that more educated are generally from wealthy families and therefore don't need to send remittances; that they tend to bring their families with them; and that they are less likely to return. However, there is a counter theoretical argument: that higher skilled tend to earn more and therefore remit a higher amount than the low skilled.

For a detailed survey of remittances and motivations to remit, see Rapoport and Docquier (2006); Yang (2011). A paper closely related to our application by Docquier et al. (2012) connects immigration policy to remittance behaviour. They study the effect of 'restrictive' and 'selective' immigration policies in the destination countries using a bilateral remittances database.[2] They find that, for a given country pair, relatively more skilled migrants will send less remittances

[2]For restrictive policies they use three proxies, namely bilateral guest worker programme, proportion of refugees among migrants and proportion of females among migrants; and for selectivity they use the point system.

if the destination country uses a selective immigration policy. We, like them, also want to analyse the impact host country's migration policy has on migrant's remittance behaviour. However, our focus is on the effect of change in a country's immigration policy on remittance behaviour, which better captures the link between remittances and education/skills.

Empirical evidence on the relationship between education and remittances is not clear as well. At a micro level, using household survey data for 11 destination countries, Bollard et al. (2011) show a positive relationship between education and amount remitted while they find mixed results for the likelihood of remittances. Using GSOEP data, Dustmann and Mestres (2010) find a negative effect of education on remittances after controlling for intentions to return and household composition at destination. At a macro level, Faini (2007), Adams (2009), Niimi et al. (2010) and Kratou et al. (2024) have found a negative relationship between remittance flows and education.

Our application using LSIA data shows a small insignificant negative effect (-2.4 pp.) of the 1996 reform on probability of remittance and no effect on remittance value, overall. Furthermore, we find a statistically significant effect for the second income quartile of immigrants – in years 1999-2000 they remit 9-13 pp. more often. Therefore, contrary to the existing literature, our results show a more nuanced relationship between education and remittances, as it only effects migrants in a specific income quartile, which the existing literature has not explored.

Our estimates stress the importance of accounting for the individual-specific trends in estimation of the treatment effects as well as the importance of controlling for compositional changes. Contrary to our findings, the conditional difference-in-difference (DDM) estimates, which do not allow heterogenous trends or control for changes in composition, find a statistically significant *increase* in remittance probability due to the policy change. Additionally, standard linear difference-in-differences (D-in-D), which naturally deals with compositional changes under linear parallel trends assumption, fails to find any significant effect of the policy on remittance probability for different income quartiles.

2 Interactive Fixed Effects Model with Compositional Changes

In this section we propose an econometric set-up that can be used to evaluate the effects of the changes in policy outlined in the Introduction.

2.1 Econometric Model

Following CK23, we only pose a model for untreated outcomes, allowing the treated outcomes to be correlated both with observables and unobservables in the model. This allows selection into treatment and means that we will only be able to identify the average treatment effect on the treated (ATT). Let $D \in \{0, 1\}$ denote the treatment status.

In order to accommodate changes in Australian labour market over our sample period, we allow unobserved heterogeneity to have a different effect in different periods by including interactive fixed effects:

$$Y_{it}(0) = \xi_i + \lambda_i' F_t + Z_i' \delta_t + U_{it} \tag{1}$$

where ξ_i is the standard additive fixed effect and λ_i measures unobserved heterogeneity that affects outcome Y_{it} differently in different periods. In that fashion we also allow observed characteristics, Z_i, to have a time-varying effect. Both ξ_i and λ_i can be correlated with Z_i. Note that if the first element of Z_i is equal to 1 and $F_t = 0, \forall t$, this model boils down to a standard two-way fixed effects model.

Throughout our analysis we assume that we only have one interactive fixed effect in model (1) as this already generalises the standard two-way fixed effect model but avoids complications implied by multiple effects (i.e. more pre-periods and more instruments needed, see Section 5.1 in CK23).

Let a binary variable τ_t indicate if an observation comes from period t and $\tau = (\tau_1, \dots, \tau_T)$. We make the following assumptions:

Assumption ID. *(a) Data consists of at least 2 pre-intervention periods.*

(b) F_t *is not constant in the pre-intervention periods.*

(c) At least one of the regressors has a constant effect, i.e. one component of δ_t *is constant across* t.

(d) Selection based on (ξ, λ, Z): $E[U|\xi, \lambda, Z, D = 0, \tau] = E[U|\xi, \lambda, Z, D = 1, \tau] = 0$.

(e) $E[(\xi, \lambda) \mid D, Z, \tau] = E[(\xi, \lambda) \mid D, Z, \tau']$ *for all* τ', τ.

(f) $0 < P(\tau_t = 1|D, Z) < 1$ *for all* $t = 1, \ldots, T$.

Assumptions ID(a)-(d) are the same as those in CK23. Assumption (c) is likely to be satisfied for variables like age at migration or region of origin dummies in our application – both the effect of age and origin on remitting are due to cultural factors, which should be stable in the short-medium term. Since we have three pre-intervention periods this assumption can be tested – an extra pre-intervention period provides an additional moment restriction which facilitates an overidentification test (see Proposition 3 in CK23 for more details).

Assumptions ID(e)-(f) replace the assumption in CK23 that observations are drawn from the same joint distribution of outcome paths and covariates irrespective of the period in which they enter the sample in CK23. Assumption (e) is a relatively strong restriction – it states that, even though (ξ, λ) can be correlated with Z and the distribution of (Z, D) can change across the periods, the mean of these individual specific unobservables stays the same. Still, if the policy change leading to compositional change is based only on observables, we can plausibly argue that there should be no change in the distribution (or at least their mean) of unobservables conditional on Z. This is the case in our application as the strengthening of the immigration policy was based mainly on observed education and English proficiency. Thus, for example, we would not expect that after 1996 migrants with higher education would be on average more or less able or more motivated (unobservables) than before 1996. Thus, Assumption (e) would hold in our application.

Assumption ID(f) is an "overlap" condition and should normally be satisfied unless the policy introduces very strong selection based on covariates. In our application, even though the 1996 reform strengthens requirement for the level of education, it still allows migrants with lower level of education to come to Australia under the points-based system (cf. Table 7).

Remark 1. *Assumption ID (e) still keeps our model not nested with the conditional parallel trends model used by Sant'Anna and Xu (2026). To see that, note that using Assumption ID:*

$$\begin{aligned}
&E[Y(0)|D=1,\tau_t=1,Z]-E[Y(0)|D=1,\tau_{t-1}=1,Z]= \\
&\qquad\qquad =E[\lambda|D=1,\tau_{t-1}=1,Z](F_t-F_{t-1})+Z'(\delta_t-\delta_{t-1}) \\
&E[Y(0)|D=0,\tau_t=1,Z]-E[Y(0)|D=0,\tau_{t-1}=1,Z]= \\
&\qquad\qquad =E[\lambda|D=0,\tau_{t-1}=1,Z](F_t-F_{t-1})+Z'(\delta_t-\delta_{t-1})
\end{aligned}$$

but we do not assume $E[\lambda|D=1,\tau,Z]=E[\lambda|D=0,\tau,Z]$ *so the differences above are, in general, not equal.*[3]

2.2 Composition Change in Observables

Although our data consists of two repeated panels (pre and post), we discuss our estimator in the context of pure repeated cross-sections. The main complication in dealing with composition changes in our model comes from the fact that we do not observe the same individuals in pre- and post-intervention periods, thus having pre- and post-panels does not help identification. Also, adjusting our results to repeated panels is straightforward.

With repeated cross-sections we observe draws from a mixture of distributions of (Y,Z,D) across time periods. Let P denote the distribution of (Y,Z,D,τ) and E be the corresponding expectation. Let t^* be the first (observed) post-intervention period and recall that $\tau=(\tau_1,\ldots,\tau_T)$ indicates the period from

[3] We do not claim that our model is more general than the one in Sant'Anna and Xu (2026) as we do assume a linear interactive fixed effect model for $Y(0)$.

which an observation was drawn. Note that for the subsample with $(D = 1, \tau_t = 1 : t \geq t^*)$ we do not observe the untreated outcome $Y(0)$.

Let $Z = (X', W')'$ and W contain covariates satisfying Assumption ID(c) with $dim(Z) \equiv K_Z$ and $dim(X) \equiv K_X$. Let $\pi_t^D(Z) = P(\tau_t = 1|D, Z)$ and $\pi_t^D = P(\tau_t = 1|D)$. For any $t^* \leq t \leq T$ define:

$$\tilde{X} = \left(X', \frac{\tau_{t^*-1}Y}{\pi^0_{t^*-1}(Z)} - \frac{\tau_{t^*-2}Y}{\pi^0_{t^*-2}(Z)}\right)$$

$$\tilde{Z} = (1 - D)Z$$

$$\Delta Y_t = \frac{\tau_t Y}{\pi^0_t(Z)} - \frac{\tau_{t^*-2}Y}{\pi^0_{t^*-2}(Z)}$$

$$\theta_t = (\delta_t^*, F_t^*)'$$

where $F_t^* = \dfrac{F_t - F_{t^*-2}}{F_{t^*-1} - F_{t^*-2}}$ and δ_t^* contains the first K_X elements of $(\delta_t - \delta_{t^*-2} - F_t^*(\delta_{t^*-1} - \delta_{t^*-2}))$ (i.e. does not include zero values corresponding to W). The following assumption translates the rank condition in CK23 to our setup.

Assumption R. *The matrix $E[\tilde{Z}\tilde{X}]$ has full rank.*

In the following theorem we show that the ATT can be identified similarly to CK23 if one reweights observations by the inverse of probability of being included in the sample.

Theorem 1. *Let Assumptions ID and R hold. Then, for any positive definite symmetric weighting matrix Σ we have:*

$$\theta_t = (E[\tilde{Z}\tilde{X}]'\Sigma E[\tilde{Z}\tilde{X}])^{-1}E[\tilde{Z}\tilde{X}]'\Sigma E[\tilde{Z}\Delta Y_t] \tag{2}$$

and

$$ATT_t = E\left[\frac{\tau_t Y}{\pi_t^1} - \frac{\tau_{t^*-2}Y}{\pi^1_{t^*-2}}\middle| D = 1\right] - E\left[\left(X', \frac{\tau_{t^*-1}Y}{\pi^1_{t^*-1}(Z)} - \frac{\tau_{t^*-2}Y}{\pi^1_{t^*-2}(Z)}\right)\theta_t \middle| D = 1\right]. \tag{3}$$

Proof. Under Assumptions ID(d)-(e) we have:

$$
\begin{aligned}
&E[Y|D=0,\tau_{t^*-1}=1,Z]-E[Y|D=0,\tau_{t^*-2}=1,Z]= \\
&\qquad =E[\lambda|D=0,\tau_{t^*-2}=1,Z](F_{t^*-1}-F_{t^*-2})+Z'(\delta_{t^*-1}-\delta_{t^*-2}) \qquad (4)\\
&E[Y|D=0,\tau_t=1,Z]-E[Y|D=0,\tau_{t^*-2}=1,Z] \\
&\qquad =E[\lambda|D=0,\tau_{t^*-2}=1,Z](F_t-F_{t^*-2})+Z'(\delta_t-\delta_{t^*-2}) \qquad (5)
\end{aligned}
$$

Note that I can write the conditional moments above using the mixture expectation. For example:

$$E[Y|D=0,\tau_t=1,Z]-E[Y|D=0,\tau_{t^*-2}=1,Z]=E\left[\frac{\tau_t Y}{\pi_t^0(Z)}-\frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^0(Z)}\Big|D=0,Z\right]$$

Now substituting $E[\lambda|D=0,\tau_{t^*-2}=1,Z]$ from (4) into (5), using Assumption ID(b) and rearranging we get:

$$E[V_t|D=0,Z]=0$$

where

$$V_t=\frac{\tau_t Y}{\pi_t^0(Z)}-\frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^0(Z)}-X'\delta_t^*-F_t^*\left(\frac{\tau_{t^*-1}Y}{\pi_{t^*-1}^0(Z)}-\frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^0(Z)}\right)=\Delta Y_t-\tilde{X}\theta_t$$

This implies that the following moment condition holds:

$$E[V_t(1-D)Z]=0$$

and by the standard formula for linear GMM we obtain (2).

Lastly, note that:

$$E[Y|D=1,\tau_t=1]=E\left[\frac{\tau_t Y}{\pi_t^1}\right].$$

Thus, following the same arguments as in CK23 we obtain:

$$
\begin{aligned}
ATT_t &= E[Y_t(1)-Y_{t^*-2}(0)|D=1]-E[Y_t(0)-Y_{t^*-2}(0)|D=1] \\
&= E\left[\frac{\tau_t Y}{\pi_t^1}-\frac{\tau_{t^*-1}Y}{\pi_{t^*-1}^1}\Big|D=1\right]-E\left[E\left[\frac{\tau_t Y}{\pi_t^1(Z)}-\frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^1(Z)}\Big|D=1,Z\right]\Big|D=1\right]
\end{aligned}
$$

and now substituting

$$E\left[\frac{\tau_t Y}{\pi_t^1(Z)}-\frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^1(Z)}\Big|D=1,Z\right]=X'\delta_t^*+F_t^*\left(\frac{\tau_{t^*-1}Y}{\pi_{t^*-1}^1(Z)}-\frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^1(Z)}\right)$$

(which can be obtained by same reasoning as above) gives (3). □

Note that the expressions in (2) and (3) are almost identical to the ones in Proposition 2 of CK23, even though they do not analyse the case of compositional change. What differs is how $\tilde{X}$, dY_t and the weights in the ATT are defined, which slightly complicates inference.

2.3 Inference

In order to estimate θ_t and ATT_t we need first step estimators of $\pi_t^0(Z)$'s. We propose to estimate them by multinomial logit when the data comes from repeated cross-sections, and by a simple logit when the data comes from repeated panels. Note that in the latter case, we have clustered sampling as, for example, observations $(Y_{it^*}, Z_i, D_i, \tau_{t^*})$ and $(Y_{it^*+1}, Z_i, D_i, \tau_{t^*+1})$ are correlated, which requires more care when estimating the variance. We state the main theorem for the case of purely repeated cross-sections and discuss the adjustments needed for repeated panels in a separate subsection.

The following assumption is needed for consistency and asymptotic normality of the cross-products involved in estimation of θ_t and the multinomial logit estimates (see e.g. Ch. 9.3. in Amemiya (1986)):

Assumption INF. *(a)* $E|Y_i|^4 < \infty$

(b) $\sup_i |Z_i| < \infty$.

(c) The matrix $E[Z_i Z_i']$ *has full rank.*

(d) The empirical distribution of Z_i *converges to a distribution function.*

Denote $p = P(D_i = 1)$. Let $\mathcal{I}(\alpha)$ be the information matrix from estimation of the multinomial logit. Define $Z_i^L = (1, Z_i')'$ and similarly let $\boldsymbol{Z}^L$ be $\boldsymbol{Z}$ appended with column of ones. The scores are:

$$s_i^D(\alpha^D) = \mathcal{I}(\alpha^D)^{-1} \sum_{s=1}^{T} (\tau_{is} - \pi_s^D(Z_i, \alpha^D)) Z_i^L$$

for $D \in \{0, 1\}$. Let the unit vector $e_s = (0, \ldots, 1, \ldots, 0)'$, $s = 1, \ldots, T$, indicate time-period s. Define:

$$P^D_{s,i} = \frac{1}{\pi^D_s(Z_i,\alpha^D)}(e_s - [\pi^D_1(Z_i,\alpha^D),\ldots,\pi^D_T(Z_i,\alpha^D)]') \otimes Z_i^{L\prime}$$

$$J = \begin{pmatrix} \mathbf{0}_{K_Z \times T(K_Z+1)} \\ E[\tilde{Z}_i \otimes (Y_i\tau_{it^*-2}P^D_{t^*-2,i} - Y_i\tau_{it^*-1}P^D_{t^*-1,i})] \end{pmatrix}$$

Further, let:

$$\tilde{V}_i = \frac{\tau_{it}Y_i}{\pi^1_t} - \frac{\tau_{it^*-2}Y_i}{\pi^1_{t^*-2}} - \left(X_i', \frac{\tau_{t^*-1}Y}{\pi^1_{t^*-1}(Z_i,\alpha^1)} - \frac{\tau_{t^*-2}Y}{\pi^1_{t^*-2}(Z_i,\alpha^1)}\right)\theta_t$$

$$M = (E[\tilde{Z}_i\tilde{X}_i]'\Sigma E[\tilde{Z}_i\tilde{X}_i]')^{-1}E[\tilde{Z}_i\tilde{X}_i]'\Sigma$$

$$S_t = ME[\tilde{Z}_i \otimes (Y_i\tau_{it^*-2}P^0_{t^*-2,i} - Y_i\tau_{it}P^0_{t,i})] + (E[\Delta Y_{it}\tilde{Z}_i] \otimes I_{K_Z})\dot{M}$$

$$\dot{M} = \left[(I_{K_Z} - E[\tilde{Z}_i\tilde{X}_i]'M)'\Sigma \otimes (E[\tilde{Z}_i\tilde{X}_i]'\Sigma E[\tilde{Z}_i\tilde{X}_i])^{-1}C - M' \otimes M\right]J$$

$$A_\theta = -E\left[D_i\left(X_i', \frac{\tau_{t^*-1}Y}{\pi^1_{t^*-1}(Z_i,\alpha^1)} - \frac{\tau_{t^*-2}Y}{\pi^1_{t^*-2}(Z_i,\alpha^1)}\right)'\right]$$

$$A_t = \begin{pmatrix} -ATT_t \\ -E[D_i\tau_{it}Y_i/(\pi^1_t)^2]/p \\ E[D_i\tau_{it^*-2}Y_i/(\pi^1_{t^*-2})^2]/p \\ \theta_{t,K_X+1}E[D_i(Y_i\tau_{it^*-1}P^1_{t^*-1,i} - Y_i\tau_{it^*-2}P^1_{t^*-2,i})]') \end{pmatrix}$$

where C is a $K_Z(K_X+1) \times K_Z(K_X+1)$ commutation matrix that satisfies $\mathrm{d}vec(\tilde{X}')/\mathrm{d}\alpha = C\mathrm{d}vec(\tilde{X})/\mathrm{d}\alpha$.

Theorem 2. *Under Assumptions ID, R and INF we have for $t \geq t^*$:*

$$\sqrt{N}(\hat{\theta}_t - \theta_t) \to N(0, Var_{\theta_t}), \qquad \sqrt{N}(\widehat{ATT}_t - ATT_t) \to N(0, Var_t),$$

where

$$Var_{\theta_t} = Var(\phi_{1i} + \phi_{2i}),$$

$$Var_t = \frac{1}{p^2}(Var(\psi_{1i} + \psi_{2i}) + Var(\psi_{3i} + \psi_{4i})),$$

$$\phi_{1i} = M\tilde{Z}_iV_i, \phi_{2i} = S_ts^0_i(\alpha^0),$$

$$\psi_{1i} = D_i\tilde{V}_i, \psi_{2i} = A_t'\begin{pmatrix} D_i - p \\ D_i\tau_{it} - P(D_i = 1, \tau_{it} = 1) \\ D_i\tau_{it^*-2} - P(D_i = 1, \tau_{it^*-2} = 1) \\ s^1_i(\alpha^1) \end{pmatrix}, \psi_{3i} = A_\theta'\phi_{1i}, \psi_{4i} = A_\theta'\phi_{2i}.$$

The independence of some of the variance components results from different parameters being estimated on treated or control samples (e.g. ATT_t vs θ_t).

We can estimate the variance by replacing all quantities above with respective sample estimators. For example, if $\hat{\alpha}^0$ denotes the coeficient estimator obtained by running multinomial logit on the control sample, $\widehat{\Delta P}_t^0$ stacks estimators of $Y_i \tau_{it^*-2} P^0_{t^*-2,i} - Y_i \tau_{it} P^0_{t,i}$ and $\hat{p}$ is the fraction of treated in the sample, we can estimate:

$$\begin{aligned}
\hat{M} &= N(\tilde{\mathbf{X}}'\tilde{\boldsymbol{Z}}\Sigma\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{Z}}\Sigma \qquad (6)\\
\hat{J} &= \begin{pmatrix} \mathbf{0}_{K_Z \times T(K_Z+1)} \\ \tilde{\boldsymbol{Z}}'\widehat{\Delta P}^0_{t^*-1} \end{pmatrix}\\
\hat{\hat{M}} &= (((I_{K_Z} - \tilde{\boldsymbol{Z}}'\boldsymbol{X}\hat{M})'\Sigma \otimes (\tilde{\mathbf{X}}'\tilde{\boldsymbol{Z}}\Sigma\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}})^{-1})C - \hat{M}' \otimes \hat{M})\hat{J}\\
\hat{S}_t &= \hat{M}\tilde{\boldsymbol{Z}}'\widehat{\Delta P}^0_t + (\widehat{\boldsymbol{\Delta Y}}_t{}'\tilde{\boldsymbol{Z}} \otimes I_{K_X+1})\hat{\hat{M}}\\
\hat{A}'_\theta &= -\frac{1}{N}\sum_{i=1}^{N} D_i \left(X'_i, \frac{\tau_{t^*-1}Y_i}{\pi^1_{t^*-1}(Z_i,\hat{\alpha}^1)} - \frac{\tau_{t^*-2}Y_i}{\pi^1_{t^*-2}(Z_i,\hat{\alpha}^1)} \right)\\
\hat{V}_i &= \frac{\tau_{it}Y_i}{\pi^0_t(Z_i,\hat{\alpha}^0)} - \frac{\tau_{it^*-2}Y_i}{\pi^0_{t^*-2}(Z_i,\hat{\alpha}^0)} - \left(X'_i, \left(\frac{\tau_{t^*-1}Y_i}{\pi^0_{t^*-1}(Z_i,\hat{\alpha}^0)} - \frac{\tau_{t^*-2}Y_i}{\pi^0_{t^*-2}(Z_i,\hat{\alpha}^0)} \right) \right)' \hat{\theta}_t
\end{aligned}$$

which gives the following estimator of the second component of the asymptotic variance of $\widehat{ATT}_t$:

$$\widehat{Var}(\psi_3 + \psi_4) = \frac{1}{(N\hat{p})^2}\hat{A}'_\theta \begin{pmatrix} \hat{M} & \hat{S}_t \end{pmatrix} \sum_{i=1}^{N} \begin{pmatrix} \tilde{Z}_i \hat{V}_i \\ s^0_i(\hat{\alpha}^0) \end{pmatrix} \begin{pmatrix} \tilde{Z}_i \hat{V}_i \\ s^0_i(\hat{\alpha}^0) \end{pmatrix}' \begin{pmatrix} \hat{M} & \hat{S}_t \end{pmatrix}' \hat{A}_\theta.$$

2.3.1 Repeated Panels

Two modifications are needed for this case. Firstly, now we just need to estimate a logit for post- and pre-periods and the predicted probabilities become $\hat{\pi}^D_t(z_i, \hat{\alpha}^D) = \Lambda(z_i, \hat{\alpha}^D)/s$, where Λ is the logistic CDF and s is the number of pre- or post-periods in the panel depending on whether $t < t^*$ or $t \geq t^*$, respectively.[4] Secondly, the variance estimator has to adjust for clustering.

[4]We assume balanced panel here. With unbalanced panel replace $1/s$ with probability of being included in a specific wave t conditional on being in pre-/post-panel.

Let now $i = 1, \ldots, N$ denote individuals in the panels, so summing over i will mean summing over individuals, not over all observations in the sample (NT). For each individual we have either T_{pre} or T_{post} observations so clustering is at the i level. Let $\hat{\psi}_{3ir}, \hat{\psi}_{4ir}$ denote the estimates of ψ_{3ir}, ψ_{4ir}, where r indexes repeated observations for respondent i. The clustering-adjusted estimator of the second component of the asymptotic variance of $\widehat{ATT}_t$ becomes now:

$$\widehat{Var}^{rp}(\psi_3 + \psi_4) = \frac{1}{(N\hat{p})^2} \sum_{i=1}^{N} \left(\sum_{r \in \mathcal{T}_i} (\hat{\psi}_{3ir} + \hat{\psi}_{4ir}) \right)^2$$

where $\mathcal{T}_i$ denotes the set of time periods when i is observed. One can adjust estimators of the other component similarly (see e.g. MacKinnon et al. (2023)).[5]

2.3.2 Monte Carlo Simulations

We perform 1000 Monte Carlo simulations to compare performance of the clustered variance estimator with cluster bootstrap in the repeated panel setting in the previous section, the case in our empirical application. For cluster bootstrap we resample at the individual level to preserve dependence within units and use quantiles of the bootstrap distribution to build confidence intervals, i.e. percentile bootstrap.

We adopt the Monte Carlo design in CK23, with one modification: we introduce an additional observed covariate X_i that enters untreated potential outcomes through a time-varying component $\delta_t X_i$. This term is included in both potential outcomes, so the true ATT remains zero in all designs as in their paper.

The untreated potential outcome is generated as

$$Y_{it}(0) = 0.1(t-1) + \xi_i + \lambda_i F_t + \beta W_i + \delta_t X_i + U_{it},$$

for $t = 1, \ldots, 5$, where $t^* = 5$ and treatment is assigned at the individual level with $P(D_i = 1) = 0.5$. We take $W_i \mid D_i \sim N(0,1)$, $\xi_i \mid D_i, W_i \sim N(D_i, 0.1)$, and

$$\lambda_i = 1 + 2D_i + \rho W_i + \zeta_i, \qquad \zeta_i \mid D_i, W_i \sim N(0, 0.1).$$

[5]One can also add degrees of freedom corrections to the estimator. With short pre- and post-panel and relatively large N these corrections will change very little and we do not pursue them.

The idiosyncratic shock is drawn as $U_{it} \mid D_i, W_i, \xi_i, \lambda_i \sim N(0, 0.1)$. The additional covariate is generated as $X_i \sim N(0, 1)$, and $\delta_t = (0, 0.2, 0.4, 0.6, 0.8)$ respectively. We set $\beta = 1$, $F_t = t$ for $t = 1, \ldots, 4$, and vary $F_5 \in \{4, 5, 8\}$ to obtain designs where the data-generating process is close to parallel trends, close to linear trends, or severely misspecified under both. The sample is constructed as a four-period pre-treatment panel for $t = 1, \ldots, 4$ together with an independent cross-section for $t = 5$. Results are in Table 10 and show that our estimator of the asymptotic variance works well and produces coverage probabilities close to nominal levels as the sample size increases. Also, the performance of clustered standard errors and cluster bootstrap is similar, with the asymptotic variance estimator showing slightly stronger undercoverage in smaller samples (especially with $F_5 = 8$). Thus, we can recommend using both methods in practice, and comparing their results.[6]

3 Immigration and Remittances

3.1 Background of Australian immigration policy

To understand the development of policies leading to the changes in the mid-1990s, one needs to put some context to Australia's immigration policy. In 1972 Australia formally ended a migration policy based on ethnicity ('white Australia policy'), replacing it with a focus on economic conditions to a limited number of workers in occupations where there was unmet demand. Eliminating racial discrimination from immigration selection resulted in increasing numbers of applicants and refugees from non-European countries and consequently higher stocks of immigrants with non-English speaking background (NESB). As an example Australia has taken refugees from conflicts in Chile (1973), Northern Cyprus (1974), Lebanon (1976-1983), Vietnam (1976-1982), Thailand (1976), East Timor (1977), Sri Lanka and El Salvador (1983) and the former Yugoslavia (1994).

[6] Note that CK23 propose a multiplier bootstrap based on the estimated influence function for their main case of (non-repeated) panel data but, intuitively, it seems that a non-parametric bootstrap at the individual level would work in their setting as well. Our MC results seem to confirm that intuition.

Until the policy change introduced in 1996 and analysed in this paper, two major trends had characterised immigration policy in Australia. First, the development of a systematic approach to immigration selection based on current labour market conditions. This took place during the late 1970s and throughout the 1980s with the introduction of the Numerical Multifactor Assessment System (NUMAS) during 1979-1982, which selected immigrants on the basis of family ties, occupational and language skills, and the points test (since 1988) which was used to set annually a minimum pass mark to be eligible for migration based on the skill level (qualifications and work experience), age and English language proficiency. Points could be gained if the applicant was qualified to work in one of the occupations listed in a Priority Occupation List, which summarises employers' views and recruitment difficulties. Pre-migration English-language testing for particular occupations (chosen annually) was also introduced in 1989, first in the medical and nursing professions and then extending to 114 professions between 1991 and 1993 (Hawthorne (2005)), with points provided for both speaking and writing language abilities. Since 1988 the migration program is divided into three streams - family reunification, skill and humanitarian - with the skill stream contributing about one third of all migrants to Australia until 1995-96.

The second trend has been the development of policies to favour the settlement and participation of migrants, especially if from NESB, to Australia's economic activity. This was accompanied by the introduction of instruments, including ad hoc data collections, to analyse their economic outcomes. Even with higher skill levels than comparable natives or migrants with an English-speaking background (Watson (1996)), NESB immigrants were characterised by substantially lower economic outcomes. To overcome a linguistic disadvantage, Australia had put in place a publicly funded system to provide new adult NESB immigrants with free language courses (as well as locally-funded technical training). Immigrants were paid to attend these courses, which lasted between one and six months and led them to attain a level of language ability that was adequate for employment. By 1990, Australia's Adult Migrant English Program (AMEP) "was the largest government-funded English teaching program for migrants worldwide, catering to over 70,000 immigrants per year including large number of unemployed professionals" (Hawthorne (2005)). The government actively pursued the private sector to adopt Equal Opportunity principles towards NESB as well as facilitated to ease the admission of 'professionals' from NESB (managed by professional associations) with the funding of specialist labour market programs designed to prepare NESB professionals for mandatory entry exams in a range of professions like medicine, engineering and nursing.

In 1996 a newly elected government introduced a number of significant changes to the migration policy, affecting the skill and family reunification but not the humanitarian streams. This new policy: (1) Abolished the social security benefit to new immigrants in the first two years after their arrival, as well as access to the Adult Migrant English Program (whose costs were now to be met by the immigrant) and labour market programs (whose costs were to be repaid after securing work); (2) Allocated the highest points weighting to employability factors, namely occupational skills, education, age, and English language ability. Age-related points for applicants over the age of 45 were abolished while bonus points were awarded to those with relevant Australian or international professional work experience, a job offer, a spouse meeting the skill application criteria, an Australian sponsor who had to provide a guarantee, and carrying A$100,000 or more in capital. By 2001 most migrants to Australia were in the skill stream; (3) Introduced additional points for occupations in demand in addition to degree-level specific (as opposed to generic) qualifications, and bonus points for qualifications obtained recently in Australia; (4) Pre-migration qualification screening was effectively outsourced to professional bodies, which could now disqualify NESB applicants from eligibility for skill migration.

3.2 Data

The Longitudinal Survey of Immigrants to Australia (LSIA) was commissioned in the early 1990s to fulfil the need to have better information on the settlement of new migrants than those available through censuses. It is based on a representative sample of 5% of migrants/refugees and comprises three surveys over a period of almost 10 years. We focus only on the first two cohorts, LSIA1 and LSIA2, as the third cohort was not surveyed in detail, limiting the comparison and use of the database. Figure 1 shows the timespan of our data diagrammatically.

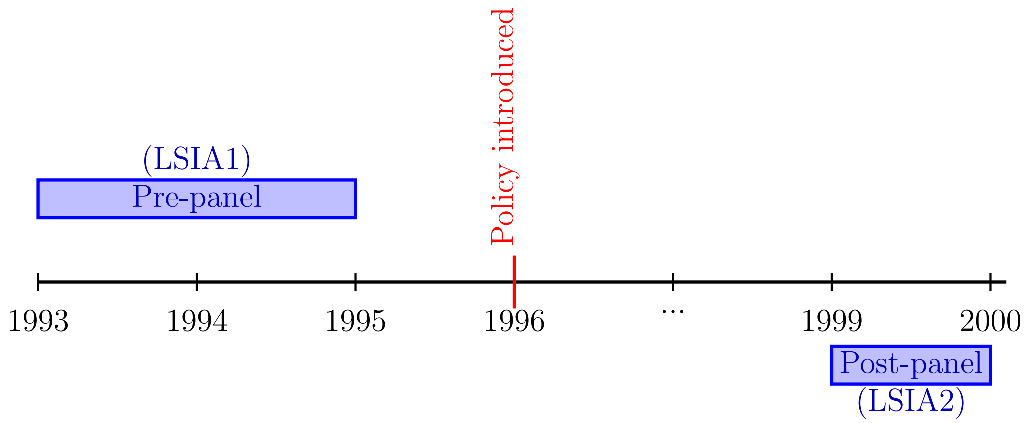


Figure 1: Longitudinal Survey of Immigrants to Australia (LSIA)

Tables 1-6 summarise some key characteristics of primary applicants, by cohort. Relative to cohort 1, there is an increase in percentage admitted on Business and Independent visas in cohort 2, and a reduction in Family and Humanitarian streams (Table 1). Migrants in the second cohort originate from a wider variety of countries relative to the first cohort, with increased arrivals from Oceania, New Zealand and the Pacific Islands, North East Asia (which includes China and Hong Kong), and Southern and Central Asia (Table 2). The second cohort had in general better education levels than the first cohort, especially with higher university degrees at the time of arrival (Table 3), and lower proportions of migrants that are not in the labour force immediately after resettlement or undertaking further education in Australia (Table 4).

Tables 5 and 6 present summary statistics on remittance behaviour based on values deflated and expressed in year 2000 AUD. The percentage of migrants remitting is smaller in cohort 2 both upon arrival (4.2% remitting compared to 7.8% in cohort 1 - Table 5) and in the subsequent 12 months (13.2% vs. 22%).

The migrants of the second cohort also appear to remit lower amounts, regardless of the length of stay in Australia (Table 6).

Table 7 reports the unconditional mean and standard deviation for the working sample, by visa type, before and after tighter immigration criteria were introduced. By focusing momentarily on average values between affected and not-affected, the two groups are different with respect to several demographic characteristics (gender, marital status), education (the affected are on average better educated) and countries of origin (the affected come from a wider group of countries), highlighting the different motivations for migration.[7] Those on Independent and Concessional visa (the affected group) are overwhelmingly admitted through the point system, and are therefore economic migrants with high prospects of immediate employability but limited or no host country support from family, employers or local institutions. In contrast, those on Preferential and Humanitarian visas, as well as employer-sponsored migrants, comprise a more heterogeneous group of settlers with a high incidence of family reunification.

Given that the data combines observations from two panels that are 4 to 7 years apart, we do not expect that the parallel trends assumption would be satisfied in the data, but rather expect that the effect of both observed and unobserved characteristics would be different in 1993-1995 than in 1999-2000. Additionally, the 1996 Australian migration policy change consists of more stringent entry requirements including higher skill level, work experience and higher English attainment, which modifies the composition of the treatment group in terms of observed characteristics (as seen above). These features motivate the use of our interactive fixed effects model with compositional changes.

[7] The affected and not-affected groups are broadly similar with respect to the country of origin (about half are born in an English-speaking country and member state of the Commonwealth) and post-migration residential choices, with three quarters of migrants settling in or around Australia's two main cities, Sydney and Melbourne. They are also similar with respect to average age (early 30s) and English language skills (almost all respondents use English in the interview). Other demographic characteristics differ significantly between the two groups. The not-affected are characterised by a higher number of women and married individuals.

3.3 Results

We apply the estimator in Section 2.3 to evaluate the effect of this policy change on immigrants' remittance behaviour by comparing the extensive (probability) and intensive (remittance value) margins of remittances for the cohort that entered after the policy change with the one that entered before the change. In order to investigate the intensive margin we define the dependent variable as a log of (deflated) remittance value for those who remit positive amounts, i.e. excluding non-remitters. As we have two post-intervention periods we can estimate the treatment effect both for year 1999 and 2000.

In Table 8 we compare our results to a simple D-in-D, the D-in-D with matching (DDM) with bootstrapped standard errors[8] and the interactive fixed effects estimator of CK23 which ignores compositional changes (IFE). Strikingly, D-in-D, DDM and IFE completely disagree about the predicted effect of the policy on remittance probability, predicting no effect, significant positive and significant negative effect, respectively.

In order to apply our estimator, we have to choose a regressor whose effect is constant across time (the instrument W). We consider three choices: age at migration, region of birth dummies or both. The regions are: Europe and former USSR, Middle East & North Africa, Southeast Asia & Oceania, Northeast Asia, Southern and Central Asia, Americas, Africa (excl. North Africa). We test the validity of these instruments using the pre-treatment panel and the procedure in Proposition 3 in CK23 and obtain p-values ranging from 0.76 to 0.99 (depending on outcome and instrument used), which means we cannot reject that instru-

[8]Bootstrapping is the prevalent method of obtaining standard errors in practice. We have also estimated the ATT for the remittance probability using the doubly robust estimator of Sant'Anna and Zhao (2020), as implemented in the DRDID package in STATA, and obtained an estimate of 0.032, sign. at 5%, which confirms our results.

ments have a constant effect.[9] In our dataset using clustered standard errors and clustered bootstrap for inference provides similar results.[10] Thus, we report confidence intervals using our developed clustered variance estimator in the main text and relegate results with cluster bootstrap to Appendix A.2.

The results from our interactive fixed effects model with compositional changes show a small insignificant negative effect of the policy on remittance behaviour. This is at odds with the estimate from the conditional DDM model, which suggests a significantly positive effect of the immigration policy on the probability of remitting money to the home country, and the original IFE model without compositional changes, which predicts a significant negative effect. Thus, we conclude that accounting for compositional changes is important for policy evaluation, at least in our application. Comparing our results with D-in-D estimates suggests that doing that *via* our method or a simple linear model under parallel trends leads to similar results, though.

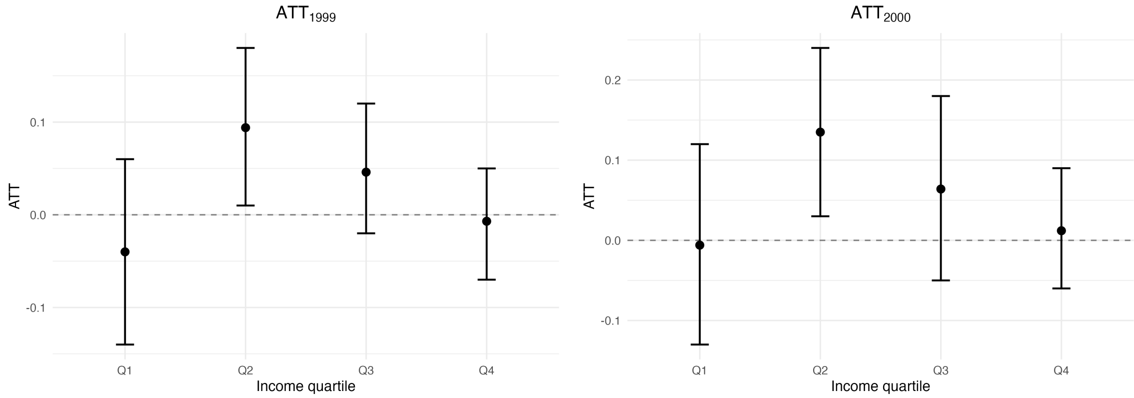
The results also confirm that using region of birth as an instrument leads to much more precise estimates than using the age at migration. The estimates with both of these instruments included are also very similar to the ones with using $W =$ *region of birth* only. Thus, we use region of birth for the rest of our analysis.

The 1996 policy reform restricted access to social security, which is expected to affect mainly the poorest immigrants. Thus, in Figure 2 we investigate the policy effect on remittance probability for different quartiles of the income distribution, where income is defined as the reported income at the point of the survey. In order to perform this exercise, we re-estimate our model on subsamples defined by income quartiles. The results show a 4 pp. drop in probability of remittance in 1999 for the first quartile and basically a zero effect in 2000, with none of these

[9]Note that with panel data this test can be implemented using standard IV commands as the estimator is technically an IV estimator with $Y_{t^*-1} - Y_{t^*-2}$ as additional regressor "instrumented" by D. See Proposition 3 in CK23.

[10]In fact, cluster bootstrap seems to produce unreasonably large confidence intervals when we use age at migration as W and behaves more erratically on small sample sizes (when we use remittance value as the dependent variable).

Figure 2: Effect on remittance probability across income quartiles (with 95% CI)



Notes: Region of birth used as instrument. Standard errors obtained by bootstrapping individuals in sub-panels (percentile bootstrap).

effects statistically significant. The fading of the effect between 1999 and 2000 may be explained twofold. Firstly, as time flies, migrants will become more attached to the Australian labour market and less reliant on social security. Second, the social security access was withdrawn only for two years so migrants entering in 1999 would be expecting to regain access in 2001 and may restart remitting in 2000, anticipating this income flow.

Additionally, we observe a significant positive effect in 1999 and in 2000 for the second quartile, with predicted increase in probability of remittance by 9.4 pp. and 13.5 pp. respectively. This constitutes a bit of a puzzle. One possible explanation is that the change in policy increases uncertainty and supports the insurance motive for remitting more. This is visible in the effect on the second quartile. However, lack of resources among people in the first income quartile prevents them from utilising this safety motive and, further, limited social security makes them less likely to remit, which generates no or slightly negative effect in the first quartile. Our results therefore show that the decision to remit is more nuanced and may explain the contrary results obtained in the existing literature.

We have also performed a similar exercise for the remittance value but sample sizes are quite small there (as it only includes those remitting non-zero amounts). Appendix A.1 contains the results. Therein we also report the results of running simple D-in-D models for different quartiles, which, interestingly, show that these models fail to capture the significant positive effect in the second income quartile. This speaks to the importance of accounting for differential individual trends via interactive fixed effects for policy evaluation.

Overall, we find some evidence of a positive effect of the new Australian immigration policy on remittances in the lower middle part of income distribution, but no convincing evidence of the effect for the remaining income quartiles. Thus, overall effect is economically negligible and we, therefore, conclude that the policy change that brought better educated to the country meant the brain drain compensation is not fulfilled – new migrants are generally slightly less likely (though, not significantly) to send and this is not compensated by an increase in the amounts transferred.

3.3.1 Partial Composition Change

In order to get an alternative view on the role of compositional changes in shaping our results we consider a middle ground between the CK23 estimator and our adjusted estimator. Namely, we assume that the composition change is only along a dimension of a small number of discrete covariates. For example, we distinguish four subpopulations based on gender and high/low income and perform estimation in each of these subpopulations to obtain four estimates of the ATT. In other words, we postulate that the CK23 assumptions hold conditional on gender and income.

Given the subgroup treatment effects we can estimate the overall ATT by:

$$\widehat{ATT}_t = \widehat{ATT}_t(m, lowinc)\widehat{P}(m, lowinc) + \ldots + \widehat{ATT}_t(f, highinc)\widehat{P}(f, highinc)$$

where, for example, $\widehat{P}(m, lowinc)$ is the frequency of low income males in the pre-intervention sample.[11] We obtain the confidence intervals by percentile bootstrap at the individual level, accounting for estimation of $\hat{P}(\cdot, \cdot)$.

Table 7 shows that the main differences between the pre- and post-panel are in terms of gender, age, income and education. Thus, we choose to condition on

[11] We have also used frequencies in the whole sample or in the post-panel and the results are very similar.

these variables in our analysis. In order to make the sample sizes nontrivial we discretise income, age and education to binary values. High education corresponds to university degree or higher and high income means income above median. Age is split into below and above 30 years old.

The results are contained in Table 9 and, generally, show around 5 pp. drop in remittance probability both in 1999 and 2000 and no significant effect of remittance value. The estimates are largely consistent across different subpopulation breakdowns.

Recall that the standard IFE model produced around a 7 pp. predicted drop in remittance probability but, once we control for composition change, this effect significantly weakens. The results in this section support the claim that controlling for composition change lowers the magnitude of the estimated effect. Our procedure only accounts for compositional change in 2×2 cells, thus partially controls for composition change, and this leads to policy effects lying between the standard IFE and our composition-change proof IFEC. This reaffirms the confidence in our main results.

4 Conclusions

Motivated by our empirical problem of evaluating the effects of the 1996 immigration policy reform in Australia, we propose a new econometrics approach that tackles the challenges of measuring the policy effect using longitudinal data from two cohorts of migrants who entered before and after new migration conditions were put in place. We extended the literature on evaluation of treatment effects with repeated panels or repeated cross-sections in the presence of interactive fixed effects by allowing compositional changes. We show that the classical conditional difference-in-difference is biased and the presence of heterogeneous trends and the compositional changes in covariates need to be taken into account in order to provide reliable estimates.

Using the Longitudinal Survey of Immigrants to Australia, we studied the impact of a change in migration policy on migrants remittance behaviour. The new policy, which was announced in 1996, dramatically changed the conditions of entry making them more stringent both in terms of initial requirements (relatively higher skills and English language proficiency) and in terms of curbing the level of support offered upon entry. Since the policy change resulted in relatively higher skilled entering the country in the second cohort, it is likely that it could have had a negative effect on the extensive and intensive margins of remittances as more educated tend to remit less, possibly because they tend to bring their families with them and/or are from wealthy families. On the other hand, education can also have a positive effect on remittance flows as migration decision might be linked to obtain some target savings to start business in the home country, especially in the presence of credit constraints, or it could be to repay the loan needed to acquire higher skills. Since the extant literature provides conflicting results, one of our contributions is to try to reconcile some of those results.

Our results point towards the direction of a slightly negative, but insignificant, relationship between stringent entry policy and the level of remittance flows. However, we find that the policy significantly increased the remittance probability of immigrants in the second income quartile, with the positive effect sustained over time. This effect is, however, compensated by a drop in remittances for the rest of the migrant population, especially among poorest immigrants. Thus, overall the reform brought better educated to the country and did not affect remittance flows of the migrants still migrating to Australia after the reform.

References

Abadie, A. (2005): "Semiparametric difference-in-differences estimators," *The Review of Economic Studies*, 72, 1–19.

Abadie, A., A. Diamond, and J. Hainmueller (2012): "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*.

——— (2015): "Comparative politics and the synthetic control method," *American Journal of Political Science*, 59, 495–510.

Adams, R. H. (2009): "The Determinants of International Remittances in Developing Countries," *World Development*, 37, 93–103.

Amemiya, T. (1986): *Advanced econometrics*, Oxford: Blackwell.

Amuedo-Dorantes, C. and S. Pozo (2006): "Remittances as insurance: evidence from Mexican immigrants," *Journal of Population Economics*, 19, 227–254.

Bai, J. (2009): "Panel data models with interactive fixed effects," *Econometrica*, 77, 1229–1279.

Batista, C. and J. Umblijs (2016): "Do migrants send remittances as a way of self-insurance?" *Oxford Economic Papers*, 68, 108–130.

Bollard, A., D. McKenzie, M. Morten, and H. Rapoport (2011): "Remittances and the brain drain revisited: The microdata show that more educated migrants remit more," *The World Bank Economic Review*, 25, 132–156.

Callaway, B. and S. Karami (2023): "Treatment effects in interactive fixed effects models with a small number of time periods," *Journal of Econometrics*, 233, 184–208.

Docquier, F., H. Rapoport, and S. Salomone (2012): "Remittances, migrants' education and immigration policy: Theory and evidence from bilateral data," *Regional Science and Urban Economics*, 42, 817–828.

Dustmann, C. and J. Mestres (2010): "Remittances and temporary migration," *Journal of Development Economics*, 92, 62–70.

Faini, R. (2007): "Remittances and the Brain Drain: Do more skilled migrants remit more?" *The World Bank Economic Review*, 21, 177–191.

Gobillon, L. and T. Magnac (2016): "Regional policy evaluation: Interactive fixed effects and synthetic controls," *Review of Economics and Statistics*, 98, 535–551.

Hawthorne, L. (2005): "Picking Winners: The recent transformation of Australia's skilled migration policy," *International Migration Review*, 39, 663–696.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997): "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme," *The Review of Economic Studies*, 64, 605–654.

Hong, S.-H. (2013): "Measuring the effect of Napster on recorded music sales: difference-in-differences estimates under compositional changes," *Journal of Applied Econometrics*, 28, 297–324.

Kratou, H., R. Pillai, and T. Sharif (2024): "Do low-skilled migrant remittances help achieve SDG 10?" *Journal of Chinese Economic and Business Studies*, 22, 31–60.

MacKinnon, J. G., M. Ø. Nielsen, and M. D. Webb (2023): "Cluster-robust inference: A guide to empirical practice," *Journal of Econometrics*, 232, 272–299.

Niimi, Y., C. Ozden, and M. Schiff (2010): "Remittances and the brain drain: skilled migrants do remit less," *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, 123–141.

Ouyang, M. and Y. Peng (2015): "The treatment-effect estimation: A case study of the 2008 economic stimulus package of China," *Journal of Econometrics*, 188, 545–557.

Piracha, M. and Y. Zhu (2012): "Precautionary savings by natives and immigrants in Germany," *Applied Economics*, 44, 2767–2776.

Rapoport, H. and F. Docquier (2006): "The economics of migrants' remittances," *Handbook of the economics of giving, altruism and reciprocity*, 2, 1135–1198.

Sant'Anna, P. H. and Q. Xu (2026): "Difference-in-Differences with compositional changes," *Journal of Econometrics*, 253, 106147.

Sant'Anna, P. H. and J. Zhao (2020): "Doubly robust difference-in-differences estimators," *Journal of Econometrics*, 219, 101–122.

Watson, I. (1996): *Opening the glass door: overseas-born managers in Australia*, Australian Government Publishing Service.

Yang, D. (2011): "Migrant remittances," *Journal of Economic Perspectives*, 25, 129–52.

Table 1: Migrants by visa categories

Major Visa Category	**LSIA 1**	**LSIA2**
Business (%)	3.44	5.79
Family (%)	65.33	62.18
Humanitarian (%)	14.12	7.44
Independent (%)	17.12	24.59

Table 2: Composition of immigrant population for cohort 1 (1993-1995) and cohort 2 (1999-2000); by region of birth

	Migrant population arrived with cohort 1		Migrant population arrived with cohort 1	
Region of Birth	Freq.	Percent	Freq.	Percent
Oceania	1,858	2.48	959	2.96
Great Britain	11,454	15.27	4,752	14.66
South Europe	6,530	8.71	2,293	7.07
Western Europe	2,918	3.89	1,187	3.66
Eastern Europe	3,505	4.67	993	3.06
Middle East	7,383	9.84	3,356	10.35
South East Asia	16,305	21.74	5,879	18.14
North East Asia	10,245	13.66	5,259	16.23
South Asia	7,208	9.61	3,437	10.6
North America	2,685	3.58	1,112	3.43
Central, South America	1,335	1.78	423	1.3
Africa	3,563	4.75	2,764	8.53
Total	**74,990**	**100**	**32,415**	**100**

Table 3: Highest level of education completed by arrival cohort (percent of the immigrant population)

Highest level of qualification completed	**Cohort 1 Wave 1**	**Cohort 1 Wave2**	**Cohort 1 Wave 3**	**Cohort 2 Wave 1**	**Cohort2 Wave 2**
Primary school	4.98	4.93	4.68	3.67	3.95
Secondary school	33.22	33.72	28.67	25.56	26.08
Trade	7.28	7.16	8.12	6.7	6.97
technical/professional	21.55	21.46	22.48	20.64	20.69
Undergraduate Degree	20.98	20.8	20.11	23.84	21.82
Post graduate degree	4.95	4.88	6.27	5.44	6.76
Higher degree	7.04	7.05	9.66	14.16	13.73
Total	**100**	**100**	**100**	**100**	**100**

Table 4: Labour Force Status in Australia by wave of interview and cohort (population)

	Cohort 1			Cohort 2	
	wave 1	**wave 2**	**wave 3**	**wave 1**	**wave 2**
Business Owner (employing others, self employed)	**2.97%**	**4.93%**	**6.40%**	**5.05%**	**8.42%**
Business owner, self employed	2.28%	3.54%	4.55%	3.69%	6.41%
Business owner employing others	0.68%	1.38%	1.86%	1.36%	2.02%
Wage earner	**31.83%**	**42.84%**	**48.44%**	**45.80%**	**53.55%**
Other employed	**0.48%**	**0.84%**	**0.13%**	**0.29%**	**0.26%**
Unemployed looking for full time or part time job	**22.63%**	**13.94%**	**10.33%**	**11.18%**	**7.20%**
Unemployed looking for full time job	20.43%	12.22%	8.39%	9.28%	5.53%
Unemployed looking for part time job	2.20%	1.72%	1.94%	1.89%	1.66%
Student	**16.21%**	**13.57%**	**6.74%**	**15.15%**	**8.09%**
Not in the labour force	**25.87%**	**23.88%**	**27.95%**	**22.53%**	**22.48%**

Table 5: Percent of migrant population remitting by interview wave (population)

	Percent remitting (population)	**Average value of remittances in AUD (population)**
cohort 1 wave 1	7.76%	93.11
cohort 1 wave 2	22.00%	348.28
cohort 1 wave 3	31.05%	770.71
cohort 2 wave 1	4.20%	76.08
cohort 2 wave 2	13.40%	355.26

Table 6: Amount of remittances sent abroad by time since arrival, AUD 2000 (population)

Remittances ($)	**Observations**	**Mean**	**Std. Dev.**
Cohort 1			
6 mths or less	4379	85.88	566.86
6 to 12 mths	794	146.89	755.74
12 to 18 mths	3491	350.12	1727.31
18 to 24 mths	980	337.45	1154.01
Cohort 2			
6 mths or less	2329	69.62	472.09
6 to 12 mths	794	93.76	497.28
12 to 18 mths	1713	367.88	1612.56
18 to 24 mths	917	336.09	1145.83

Table 7: Means and differences: Balance tests

	Affected[a]			*Not Affected*[b]		
Variables	**Before**	**After**	***Difference***	**Before**	**After**	***Difference***
Probability remit	.146 (.354)	.107 (.309)	*-.039****	.126 (.331)	.068 (.253)	*-.058****
Amount remitted	6.950 (.938)	7.590 (.735)	*.640****	6.764 (.857)	7.396 (.870)	*.632****
Age	33.2 (6.52)	33.0 (6.63)	*-.02*	33.8 (9.85)	35.5 (10.27)	*1.7****
Female	.286 (.452)	.347 (.476)	*.061****	.503 (.500)	.497 (.500)	*-.006*
Married	.594 (.491)	.611 (.488)	*.017*	.790 (.407)	.748 (.434)	*-.042****
N household	2.58 (.653)	2.55 (.647)	*-.03**	2.59 (.570)	2.60 (.561)	*.01*
N relatives HC	5.88 (2.84)	5.77 (2.87)	*-.11*	5.24 (2.89)	4.90 (2.91)	*-.34****
N relative AU	.776 (1.37)	.617 (1.16)	*-.159****	1.34 (2.11)	1.70 (2.24)	*.36****
Previous visits	.454 (.498)	.662 (.473)	*.208****	.490 (.500)	.504 (.500)	*.014*
Education HS-	.411 (.492)	.318 (.466)	*-.093****	.682 (.466)	.654 (.476)	*-.028***
BA	.339 (.473)	.337 (.473)	*-.002*	.181 (.385)	.191 (.392)	*.010*
Postgraduate	.099 (.299)	.111 (.314)	*.012*	.048 (.213)	.046 (.210)	*-.002*
Higher	.150 (.358)	.234 (.424)	*.084****	.090 (.286)	.109 (.312)	*.019****
Interview E	.697 (.460)	.679 (.467)	*-.018*	.684 (.465)	.652 (.476)	*-.032****
Participates	.809 (.393)	.882 (.323)	*.073****	.591 (.492)	.586 (.492)	*-.005*
Income: low	.125 (.331)	.143 (.350)	*.018*	.308 (.462)	.289 (.453)	*-.019**
Medium-L	.281 (.449)	.112 (.315)	*-.169****	.313 (.464)	.238 (.426)	*-.075****
Medium-H	.281 (.450)	.190 (.393)	*-.091****	.184 (.388)	.173 (.378)	*-.011*
High	.294 (.456)	.548 (.498)	*.254****	.178 (.383)	.280 (.449)	*.102****
COB: NW Europe	.215 (.411)	.219 (.413)	*.004*	.187 (.344)	.159 (.366)	*-.028****
SE Europe	.110 (.313)	.066 (.418)	*-.044****	.168 (.373)	.195 (.396)	*.027****
MENA	.062 (.242)	.016 (.127)	*-.046****	.132 (.338)	.086 (.281)	*-.046****
SE Asia	.140 (.347)	.156 (.363)	*.016*	.229 (.420)	.169 (.375)	*-.060****
E Asia	.186 (.389)	.179 (.384)	*-.007*	.120 (.325)	.151 (.358)	*.031****
S Asia	.148 (.355)	.154 (.362)	*.006*	.043 (.202)	.044 (.205)	*.001*
N America	.012 (.111)	.015 (.124)	*.003*	.038 (.191)	.059 (.235)	*.021****
Latin America	.060 (.238)	.023 (.150)	*-.037****	.061 (.239)	.053 (.225)	*-.008*
Africa	.049 (.216)	.094 (.292)	*.045****	.047 (.212)	.057 (.233)	*.010***
Oceania	.018 (.132)	.076 (.265)	*.58****	.025 (.158)	.025 (.156)	*.0*
Gini coefficient	.397 (.088)	.400 (.096)	*.003*	.386 (.082)	.378 (.086)	*-.008****
Network	.051 (.086)	.055 (.088)	*.004****	.029 (.059)	.033 (.066)	*.004****
GDP: low	.265 (.442)	.284 (.451)	*.019*	.241 (.428)	.273 (.445)	*.032****
Medium-L	.222 (.415)	.210 (.408)	*-.012*	.261 (.439)	.226 (.418)	*-.035****
Medium-H	.187 (.390)	.200 (.400)	*.013*	.240 (.427)	.222 (.415)	*-.018**
High	.326 (.469)	.305 (.460)	*-.021*	.258 (.438)	.279 (.449)	*.021***
N	2,491	1,093		4,347	3,023	

Notes: Only the first two waves of each cohort are used for before/after comparability.

Standard deviation in parentheses. a Includes (i) Family concessional and (ii) skilled independent visa categories.

b Includes: (i) Family preferential, (ii) employer nomination and (iii) humanitarian visa categories.

Table 8: ATT estimates

	Probability of remittance		Log of remittance value	
Model	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]
D-in-D	0.004		-0.151	
	[-0.02, 0.03]		[-0.34, 0.04]	
DDM	0.048*		-0.190	
	[-0.00, 0.10]		[-0.54, 0.16]	
IFE	-0.070**	-0.072**	-1.134	-2.959**
	[-0.13, -0.01]	[-0.13, -0.02]	[-3.22, 0.86]	[-5.95, -0.09]
IFE with compositional changes (IFEC)				
W = *age at migration*	-0.058*	0.025	-7.791	-0.232
	[-0.13, 0.01]	[-0.02, 0.07]	[-22.24, 6.66]	[-4.26, 3.80]
W = *region of birth*	-0.024	0.017	1.862	1.668
	[-0.07, 0.02]	[-0.03, 0.07]	[-3.12, 6.85]	[-1.35, 4.69]
W = *both*	-0.022	0.017	1.871	1.641
	[-0.07, 0.02]	[-0.03, 0.07]	[-3.10, 6.84]	[-1.38, 4.66]
NT	9170	8685	1751	1929

Notes: ** Significant at 5% level, * Significant at 10% level. We control for: gender, education, income, marital status, labour force status, language of the interview (or subset of these depending on the model). For DDM we match on the variables: gender, marital status, age at migration, labour force status in home country, number of relatives in Australia and abroad, value of funds on arrival to AU, and include other aforementioned covariates in the D-in-D estimation. Remittance values have been adjusted for inflation. D-in-D and IFEC standard errors clustered at respondent level. 2000 bootstrap replications for obtaining confidence intervals for DDM and IFE.

Table 9: ATT estimates - partial composition change

Subgroups	Probability of remittance		Log of remittance value	
	ATT_{1999}	ATT_{2000}	ATT_{1999}	ATT_{2000}
	[95% CI]	[95% CI]	[95% CI]	[95% CI]
Gender × Income	-0.045*	-0.052*	-0.799	-2.090
	[-0.09, 0.01]	[-0.10, 0.00]	[-2.68, 1.23]	[-4.98, 0.82]
Gender × Education	-0.039	-0.051*	-1.540	-3.231**
	[-0.10, 0.01]	[-0.12, 0.01]	[-3.34, 0.40]	[-6.38, -0.19]
Age × Income	-0.052**	-0.052**	-0.986	-1.801
	[-0.10, -0.01]	[-0.11, -0.00]	[-2.93, 1.02]	[-4.63, 1.38]
NT	15432	15432	2335	2335

Notes: ** Significant at 5% level, * Significant at 10% level. W = region of birth. We control for: gender, education, income, marital status, labour force status, language of the interview (or subset of these depending on the model). Remittance values have been adjusted for inflation. Cluster percentile bootstrap CIs with 2000 bootstrap replications.

Table 10: MC simulations – coverage probabilities

		Asymptotic			Cluster bootstrap		
n	F_5	90%	95%	99%	90%	95%	99%
100	4	0.834	0.918	0.975	0.864	0.923	0.983
500	4	0.866	0.928	0.980	0.882	0.938	0.986
1000	4	0.902	0.951	0.988	0.897	0.943	0.984
100	5	0.855	0.911	0.968	0.887	0.929	0.984
500	5	0.869	0.926	0.981	0.895	0.945	0.988
1000	5	0.897	0.945	0.987	0.905	0.956	0.988
100	8	0.869	0.921	0.981	0.906	0.953	0.987
500	8	0.842	0.913	0.976	0.869	0.936	0.984
1000	8	0.892	0.944	0.988	0.900	0.956	0.989

Notes: 1000 MC simulations. 999 bootstrap replications.

A Additonal Results

A.1 ATT across Income Quartiles

Table A1: ATT estimates across income quartiles

			Probability of remittance			Log of remittance value	
Model		NT	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]	NT	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]
D-in-D	Q1	3306	0.012		321	0.059	
			[-0.04,0.06]			[-0.56, 0.68]	
	Q2	4535	0.046		558	-0.274	
			[-0.01,0.10]			[-1.13,0.59]	
	Q3	3274	-0.011		738	-0.116	
			[-0.07,0.05]			[-0.46,0.23]	
	Q4	4084	-0.008		698	0.022	
			[-0.05,0.04]			[-0.35,0.39]	
IFEC	Q1	1714/1304	-0.040	-0.006	213/221	-0.939	1.693
			[-0.14, 0.06]	[-0.13, 0.12]		[-18.62, 16.75]	[-12.09, 15.47]
	Q2	2592/2488	0.094**	0.135**	412/448	2.505	4.432
			[0.01, 0.18]	[0.03, 0.24]		[-8.91, 13.92]	[-6.32, 15.18]
	Q3	2256/2261	0.046	0.064	598/660	3.699	2.654
			[-0.02, 0.12]	[-0.05, 0.18]		[-11.81, 19.21]	[-3.63, 8.94]
	Q4	2608/2632	-0.007	0.012	528/600	0.918	1.815
			[-0.07, 0.05]	[-0.06, 0.09]		[-5.12, 6.95]	[-1.96, 5.59]

Notes: ** Significant at 5% level, * Significant at 10% level. We control for: gender, education, income, marital status, labour force status, language of the interview (or subset of these depending on the model). Remittance values have been adjusted for inflation. Region of birth used as W in the interactive fixed effect model with compositional changes (IFEC). D-in-D and IFEC standard errors clustered at respondent level. NT column reports the sample sizes used for ATT_{1999} and ATT_{2000}, respectively.

A.2 Results with Cluster Bootstrap

Table A2: ATT estimates, cluster bootstrap

Model	Probability of remittance		Log of remittance value	
	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]
D-in-D	0.004		-0.151	
	[-0.02, 0.03]		[-0.34, 0.04]	
DDM	0.048^{*}		-0.190	
	[-0.00, 0.10]		[-0.54, 0.16]	
IFE	-0.070^{**}	-0.072^{**}	-1.134	-2.959^{**}
	[-0.13, -0.01]	[-0.13, -0.02]	[-3.22, 0.86]	[-5.95, -0.09]
IFE with compositional changes (IFEC)				
W = *age at migration*	-0.058	0.025	-7.791	-0.232
	[-0.66, 0.33]	[-0.18, 0.14]	[-81.13, 57.44]	[-17.22, 18.89]
W = *region of birth*	-0.024	0.017	1.862^{**}	1.668
	[-0.07, 0.03]	[-0.04, 0.07]	[0.51, 3.20]	[-0.47, 3.69]
W = *both*	-0.022	0.017	1.871	1.641
	[-0.07, 0.03]	[-0.03, 0.07]	[-1.22, 4.47]	[-0.33, 3.70]
NT	9170	8685	1751	1929

Notes: ** Significant at 5% level, * Significant at 10% level. We control for: gender, education, income, marital status, labour force status, language of the interview (or subset of these depending on the model). For DDM we match on the variables: gender, marital status, age at migration, labour force status in home country, number of relatives in Australia and abroad, value of funds on arrival to AU, and include other aforementioned covariates in the D-in-D estimation. Remittance values have been adjusted for inflation. 2000 bootstrap replications.

Table A3: ATT estimates across income quartiles, cluster bootstrap

			Probability of remittance			Log of remittance value	
Model		NT	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]	NT	ATT_{1999} [95% CI]	ATT_{2000} [95% CI]
D-in-D	Q1	3306	0.012		321	0.059	
			[-0.04,0.06]			[-0.56, 0.68]	
	Q2	4535	0.046		558	-0.274	
			[-0.01,0.10]			[-1.13,0.59]	
	Q3	3274	-0.011		738	-0.116	
			[-0.07,0.05]			[-0.46,0.23]	
	Q4	4084	-0.008		698	0.022	
			[-0.05,0.04]			[-0.35,0.39]	
IFEC	Q1	1714/1304	-0.040	-0.006	213/221	-1.448	1.728
			[-0.16, 0.07]	[-0.16, 0.15]		[-9.52, 6.20]	[-5.84, 10.36]
	Q2	2592/2488	0.094*	0.135**	412/448	2.505	4.432*
			[-0.00, 0.21]	[0.02, 0.26]		[-1.34, 7.41]	[-0.58, 11.18]
	Q3	2256/2261	0.046	0.064	598/660	3.699**	2.654
			[-0.08, 0.17]	[-0.11, 0.21]		[1.00, 5.80]	[-2.25, 6.27]
	Q4	2608/2632	-0.007	0.012	528/600	0.918	1.815
			[-0.09, 0.07]	[-0.08, 0.09]		[-1.25, 3.06]	[-1.00, 4.59]

Notes: ** Significant at 5% level, * Significant at 10% level. We control for: gender, education, income, marital status, labour force status, language of the interview (or subset of these depending on the model). Remittance values have been adjusted for inflation. Region of birth used as W in the interactive fixed effect model with compositional changes (IFEC). 2000 bootstrap replications.

B Proof of Theorem 2

B.1 Asymptotic Variance of $\hat{\theta}_t$

Let $\hat{M}(\alpha^0)$ be a version of $\hat{M}$ in equation (6) with $\tilde{\boldsymbol{X}}$ evaluated at the true probabilities $\pi_t^0(Z_i, \alpha^0)$. Define $\widehat{\Delta \boldsymbol{Y}_t}$ to be a version of $\Delta \boldsymbol{Y}_t$ with true probabilities replaced with estimators $\hat{\pi}_t^0(\cdot)$. The estimator of θ_t can be written as:

$$
\begin{aligned}
\hat{\theta}_t &= \hat{M}(\alpha^0)\frac{1}{N}\tilde{\boldsymbol{Z}}'\widehat{\Delta \boldsymbol{Y}_t} + (\hat{M} - \hat{M}(\alpha^0))\frac{1}{N}\tilde{\boldsymbol{Z}}'\widehat{\Delta \boldsymbol{Y}_t} \\
&= \hat{M}(\alpha^0)\frac{1}{N}\tilde{\boldsymbol{Z}}'\Delta \boldsymbol{Y}_t + \hat{M}(\alpha^0)\frac{1}{N}\tilde{\boldsymbol{Z}}'(\widehat{\Delta \boldsymbol{Y}_t} - \Delta \boldsymbol{Y}_t) + (\hat{M} - \hat{M}(\alpha^0))\frac{1}{N}\tilde{\boldsymbol{Z}}'\widehat{\Delta \boldsymbol{Y}_t} \\
&= \theta_t + \underbrace{\hat{M}(\alpha^0)\frac{1}{N}\tilde{\boldsymbol{Z}}'\boldsymbol{V}}_{(1)} + \underbrace{\hat{M}(\alpha^0)\frac{1}{N}\tilde{\boldsymbol{Z}}'(\widehat{\Delta \boldsymbol{Y}_t} - \Delta \boldsymbol{Y}_t)}_{(2)} + \underbrace{(\hat{M} - \hat{M}(\alpha^0))\frac{1}{N}\tilde{\boldsymbol{Z}}'\widehat{\Delta \boldsymbol{Y}_t}}_{(3)}
\end{aligned}
$$

Now under Assumption INF we have $\hat{M}(\alpha^0) = M + O_p(N^{-1/2})$ so:

$$
(1) = M\frac{1}{N}\sum_{i=1}^{N} V_i\tilde{Z}_i + o_p(N^{-1/2}) \equiv \frac{1}{N}\sum_{i=1}^{N}\phi_{1i} + o_p(N^{-1/2}). \qquad (7)
$$

By asymptotic normality of the multinomial logit estimator we also have:

$$
\begin{aligned}
\widehat{\Delta Y_t} - \Delta Y_t &= \frac{\mathrm{d}}{\mathrm{d}\alpha^0}\left(\frac{\tau_t Y}{\pi_t^0(Z,\alpha^0)} - \frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^0(Z,\alpha^0)}\right)'(\hat{\alpha}^0 - \alpha^0) + o_p(N^{-1/2}) \\
&= \Bigg(-\frac{\tau_t Y}{(\pi_t^0(Z,\alpha^0))^2}\pi_t^0(Z,\alpha^0)(e_t - [\pi_1^0(Z,\alpha^0),\ldots,\pi_T^0(Z,\alpha^0)]')\otimes Z^{L\prime} \\
&\quad + \frac{\tau_{t^*-2}Y}{(\pi_{t^*-2}^0(Z,\alpha^0))^2}\pi_{t^*-2}^0(Z,\alpha^0)(e_{t^*-2} - [\pi_1^0(Z,\alpha^0),\ldots,\pi_T^0(Z,\alpha^0)]')\otimes Z^{L\prime}\Bigg)'(\hat{\alpha}^0 - \alpha^0) \\
&\quad + o_p(N^{-1/2}) \\
&= (Y\tau_{t^*-2}P_{t^*-2}^0 - Y\tau_t P_t^0)'(\hat{\alpha}^0 - \alpha^0) + o_p(N^{-1/2})
\end{aligned}
$$

This, CLT and asymptotic linear representation $\hat{\alpha}^0 - \alpha^0 = \frac{1}{N}\sum_{i=1}^{N} s_i^0(\alpha^0) + o_p(N^{-1/2})$ now imply:

$$
(2) = ME[\tilde{Z}_i \otimes (Y_i\tau_{it^*-2}P_{t^*-2,i}^0 - Y_i\tau_{it}P_{t,i}^0)]\frac{1}{N}\sum_{i=1}^{N} s_i^0(\alpha^0) + o_p(N^{-1/2}) \qquad (8)
$$

By consistency and asymptotic normality of multinomial logit we have $\frac{1}{N}\tilde{\boldsymbol{Z}}'\widehat{\Delta \boldsymbol{Y}_t} =$

$E[\Delta Y_{it}\tilde{Z}_i] + o_p(N^{-1/2})$, which implies:

$$
\begin{aligned}
(3) &= vec\left((\hat{M} - \hat{M}(\alpha^0))E[\Delta Y_{it}\tilde{Z}_i]\right) + o_p(N^{-1/2}) \\
&= (E[\Delta Y_{it}\tilde{Z}_i] \otimes I_{K_z})vec(\hat{M} - \hat{M}(\alpha^0)) + o_p(N^{-1/2})
\end{aligned}
$$

Furthermore:

$$
\frac{\mathrm{d}vec(\hat{M} - \hat{M}(\alpha^0))}{\mathrm{d}\alpha^0} = vec\left((\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{Z}}\Sigma\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}})^{-1}\frac{\mathrm{d}\frac{1}{N}(\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}})'}{\mathrm{d}\alpha^0}\Sigma[I_{K_Z} - \frac{1}{N}\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}}\hat{M}]\right) - vec\left(\hat{M}\frac{\mathrm{d}\frac{1}{N}\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}}}{\mathrm{d}\alpha^0}\hat{M}\right)
$$

Now observe that only the last column of $\tilde{\boldsymbol{X}}$ depends on α^0, which means only the last column of $\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}}$ depends on α^0 and using CLT we have:

$$
\frac{\mathrm{d}\frac{1}{N}\tilde{\boldsymbol{Z}}'\tilde{\boldsymbol{X}}}{\mathrm{d}\alpha^0} = \begin{pmatrix} \mathbf{0}_{K_Z \times T(K_Z+1)} \\ \frac{1}{N}\tilde{\boldsymbol{Z}}'J_\alpha \end{pmatrix} = J + o_p(N^{-1/2})
$$

where:

$$
J_\alpha = \begin{pmatrix} (Y_1\tau_{1t^*-2}P^0_{t^*-2,1} - Y_1\tau_{1t^*-1}P^0_{t^*-1,1})Z^L_{11} & \dots & (Y_1\tau_{1t^*-2}P^0_{t^*-2,1} - Y_1\tau_{1t^*-1}P^0_{t^*-1,1})Z^L_{1K_Z+1} \\ \vdots & \ddots & \vdots \\ (Y_N\tau_{Nt^*-2}P^0_{t^*-2,N} - Y_N\tau_{Nt^*-1}P^0_{t^*-1,N})Z^L_{N1} & \dots & (Y_N\tau_{Nt^*-2}P^0_{t^*-2,N} - Y_N\tau_{Nt^*-1}P^0_{t^*-1,N})Z_{NK_Z+1} \end{pmatrix}
$$

and J is defined in the text. Finally, using $vec(ABC) = (C' \otimes A)vec(B)$ and similar reasoning as above for other terms, we obtain:

$$
\begin{aligned}
\frac{\mathrm{d}vec(\hat{M} - \hat{M}(\alpha^0))}{\mathrm{d}\alpha^0} &= \left[(I_{K_Z} - E[\tilde{Z}_i\tilde{X}_i]'M)'\Sigma \otimes (E[\tilde{Z}_i\tilde{X}_i]'\Sigma E[\tilde{Z}_i\tilde{X}_i])^{-1}C - M' \otimes M\right] J + o_p(N^{-1/2}) \\
&\equiv \dot{M} + o_p(N^{-1/2})
\end{aligned}
$$

Hence, the second order Taylor expansion of $vec(\hat{M} - \hat{M}(\alpha^0))$ gives:

$$
\begin{aligned}
(3) &= (E[\Delta Y_{it}\tilde{Z}_i] \otimes I_{K_z})\dot{M}(\hat{\alpha}^0 - \alpha^0) + o_p(N^{-1/2}) \\
&= (E[\Delta Y_{it}\tilde{Z}_i] \otimes I_{K_z})\dot{M}\frac{1}{N}\sum_{i=1}^{N} s^0_i(\alpha^0) + o_p(N^{-1/2}) \qquad (9)
\end{aligned}
$$

Putting (7), (8) and (9) together gives:

$$
\begin{aligned}
&\sqrt{N}(\hat{\theta}_t - \theta_t) = \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Bigg\{MV_i\tilde{Z}_i + \Bigg(ME[\tilde{Z}_i \otimes (Y_i\tau_{it^*-2}P^0_{t^*-2,i} - Y_i\tau_{it}P^0_{t,i})] \\
&\qquad + (E[\Delta Y_{it}\tilde{Z}_i] \otimes I_{K_z})\dot{M}\Bigg)s^0_i(\alpha^0)\Bigg\} + o_p(1) \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left\{MV_i\tilde{Z}_i + S_t s^0_i(\alpha^0)\right\} + o_p(1) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\{\phi_{1i} + \phi_{2i}\} + o_p(N^{-1/2})
\end{aligned}
$$

and the asymptotic normality follows from the CLT.

B.2 Asymptotic Variance of $\widehat{ATT}_t$

Recall that $\hat{p} = \frac{1}{N}\sum_{i=1}^{N} D_i$. We can write $\widehat{ATT}_t$ as:

$$\widehat{ATT}_t = \frac{1}{\hat{p}}\frac{1}{N}\sum_{i=1}^{N} D_i \left[\frac{\tau_{it}Y_i}{\hat{\pi}_t^1} - \frac{\tau_{it^*-2}Y_i}{\hat{\pi}_{t^*-2}^1} - \left(X_i', \frac{\tau_{t^*-1}Y}{\pi_{t^*-1}^1(Z_i,\hat{\alpha}^1)} - \frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^1(Z_i,\hat{\alpha}^1)}\right)' \hat{\theta}_t\right].$$

Let $\hat{\gamma} = (\hat{\pi}_t^1, \hat{\pi}_{t^*-2}^1, \hat{\alpha}^{1\prime}, \hat{\theta}_t')'$ and define:[12]

$$f_i(\hat{\gamma}) = D_i \left[\frac{\tau_{it}Y_i}{\hat{\pi}_t^1} - \frac{\tau_{it^*-2}Y_i}{\hat{\pi}_{t^*-2}^1} - \left(X_i', \frac{\tau_{t^*-1}Y}{\pi_{t^*-1}^1(Z_i,\hat{\alpha}^1)} - \frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^1(Z_i,\hat{\alpha}^1)}\right) \hat{\theta}_t\right]$$

Consider the following decomposition:

$$\widehat{ATT}_t - ATT_t = \underbrace{\frac{1}{p}\frac{1}{N}\sum_{i=1}^{N}\{f_i(\gamma) - E[f_i(\gamma)]\}}_{①} - \underbrace{\frac{\hat{p}-p}{p\hat{p}}\frac{1}{N}\sum_{i=1}^{N} f_i(\hat{\gamma})}_{②} + \underbrace{\frac{1}{p}\frac{1}{N}\sum_{i=1}^{N}\{f_i(\hat{\gamma}) - f_i(\gamma)\}}_{③}$$

Firstly, note that:

$$① = \frac{1}{p}\frac{1}{N}\sum_{i=1}^{N} D_i\tilde{V}_i = \frac{1}{p}\frac{1}{N}\sum_{i=1}^{N}\psi_{1i}$$

Next, we have:

$$\frac{\hat{p}-p}{\hat{p}} = \frac{1}{p}\frac{1}{N}\sum_{i=1}^{N}(D_i - p) + o_p(N^{-1/2})$$

By the result in the previous section and standard arguments we have $\hat{\gamma} - \gamma = O_p(N^{-1/2})$, which gives:

$$③ = \frac{1}{p}\frac{1}{N}\sum_{i=1}^{N}\Delta f_i(\gamma)'(\hat{\gamma} - \gamma) + o_p(N^{-1/2})$$

where

$$\Delta f_i(\gamma)' = \begin{pmatrix} -\frac{D_i\tau_{it}Y_i}{(\pi_t^1)^2} \\ \frac{D_i\tau_{it^*-2}Y_i}{(\pi_{t^*-2}^1)^2} \\ \theta_{tK_X+1}D_i(Y_i\tau_{it^*-1}P_{t^*-1,i}^1 - Y_i\tau_{it^*-2}P_{t^*-2,i}^1)' \\ -D_i\left(X_i', \frac{\tau_{t^*-1}Y}{\pi_{t^*-1}^1(Z_i,\alpha^1)} - \frac{\tau_{t^*-2}Y}{\pi_{t^*-2}^1(Z_i,\alpha^1)}\right)' \end{pmatrix}$$

[12] We drop the t subscript for notational brevity.

Now using $\hat{\theta}_t - \theta_t = \frac{1}{N}\sum_{i=1}^{N}\left\{MV_i\tilde{Z}_i + S_t s_i^0(\alpha^0)\right\} + o_p(N^{-1/2})$, $\hat{\alpha}^1 - \alpha_1 = \frac{1}{N}\sum_{i=1}^{N} s_i^1(\alpha^1) + o_p(N^{-1/2})$ and grouping terms we obtain:

$$\widehat{ATT}_t - ATT_t = \frac{1}{p}\frac{1}{N}\sum_{i=1}^{N}\left(\psi_{1i} + \hat{A}_t' \begin{pmatrix} D_i - p \\ (D_i\tau_{it} - P(D_i = 1, \tau_{it} = 1))/p \\ (D_i\tau_{it^*-2} - P(D_i = 1, \tau_{it^*-2} = 1))/p \\ s_i^1(\alpha^1) \end{pmatrix} + \hat{A}_\theta' MV_i\tilde{Z}_i + \hat{A}_\theta' S_t s_i^0(\alpha^0)\right) + o_p(N^{-1/2}) \tag{10}$$

where $\hat{A}_t$ and $\hat{A}_\theta$ replace elements in A_t and A_θ with its sample equivalents.

Now in order to obtain our main result we need consistency of $\hat{A}_t$ and $\hat{A}_\theta$ for A_t and A_θ. Consistency of $\widehat{ATT}_t$ follows from consistency of the multinomial logit estimator $\hat{\alpha}^1$, consistency of $\hat{\theta}_t$ and consistency of $\hat{\pi}^1$'s, which follow from our assumptions.

Now, applying the Slutsky's theorem and the CLT we obtain asymptotic normality of $\sqrt{N}(\widehat{ATT}_t - ATT_t)$. Note that the last two terms in (10) are estimated on the control sample ($D_i = 0$) while the other elements use the subsample with $D_i = 1$. Thus, these two terms are independent of the rest.