November 2022

Crossover Design and Evaluation of Labor Market Effects of Minimum Wages in China

Zhengxiong Yang, Shiwei Zhang, Mingwei Liu

Working Paper 2022.01

Rutgers
School of Management and Labor Relations
CENTER FOR GLOBAL WORK AND EMPLOYMENT

Crossover Design and Evaluation of Labor Market Effects of Minimum Wages in China

Zhengxiong Yang [a], Shiwei Zhang [b], Mingwei Liu [c]

[a] *Nanjing University of Finance and Economics, China, yangzx@nufe.edu.cn*

[b] *Jilin University, China, zhangsw@jlu.edu.cn*

[c] *Rutgers University, the United States of America, mingwei@smlr.rutgers.edu*

Abstract

We propose a crossover design, in which one of the two study groups crosses over from the treatment group in the first period to the control group in the second period, as an extension of the Difference-in-Differences (DID) method for evaluating treatment effects. We show that the crossover design can make reliable inferences on treatment effects while relaxing the parallel trends assumption of the DID design. We apply the crossover design to assess minimum wages effects in China, of which the validity of existing DID-based findings is in question due to the lack of a test of parallel trends.

Keywords: Crossover Design; Difference-in-Differences; Minimum Wages; China

JEL classification: C21; J8

Correspondence: Mingwei Liu, Rutgers University, mingwei@smlr.rutgers.edu

1. Introduction

Difference-in-Differences (DID) is a popular method to estimate the causal effect of a treatment. The key premise of the DID design is the parallel trends assumption—the outcome variables of the treatment group and the control group would experience the same trends without the treatment. The conventional approach to support the assumption is to compare the trends of treated and controls before the intervention, which is usually referred to as a placebo test. Implementation of a placebo test requires at least two periods be observed in which both the treatment and control groups are untreated.

However, this condition is often not met when applying the DID design. Units may be treated in different periods and there may be multiple treatments on the same unit. For example, in the two groups, three-periods design, assume that neither group is treated in the first period and that the treatment is assigned to one group in the third period: $[(0,0)|(.,.)|(1,0)]$ (1 denotes treated and 0 untreated in a specific period), then there are four types of three-periods design: $[(0,0)|(0,0)|(1,0)]$, $[(0,0)|(1,0)|(1,0)]$, $[(0,0)|(0,1)|(1,0)]$, $[(0,0)|(1,1)|(1,0)]$. Among these designs, the placebo test can only be conducted for the first type, i.e., the conventional DID design. How to assess the validity of the DID estimator for other types, however, has been largely overlooked in extant literature.

Our study focuses on the third type, in which one of the two groups crosses over from the treatment group in one period to the control group in the subsequent period. We refer to this setup as a crossover design and study its identification and inference procedures. We show that the crossover design can make a credible inference on the treatment effect and has less restrictive assumptions than the DID design as it does not rely on the parallel trends. In addition, under assumptions that the treatment effects are homogeneous and the time trend differences are stable, it can not only identify the treatment effect on the treated (ATT) as the conventional DID design does, but also allow an inference on the average treatment effect (ATE) since both groups are treated. Further, the crossover design can help identify and reject spurious estimators.

We apply the crossover design to evaluate the wages and employment effects of minimum wages in China. Because Chinese local governments frequently adjust minimum wages, it is very difficult to find appropriate data to implement the DID design with a placebo test. As such, the validity of existing DID-based findings is in question. This feature of Chinese minimum wages adjustment, however, makes it convenient to implement a crossover design. Using data from the 2011-2014 "China Migrant Dynamics Survey" (CMDS hereafter), we show that the increase of minimum wages has a significant and positive effect on wages of low-skilled rural migrant workers in the central and western regions of China. The effect, however, becomes insignificant for high-skilled rural migrant workers. We find no evidence that minimum wages adversely affect employment. Instead, the increase of minimum wages has a positive effect on employment of low-skilled rural migrant workers. These results suggest a monopsonistic labor market for Chinese rural migrant workers.

Our study makes two contributions. First, we contribute to the DID literature with a new design. A recent extension of the DID method relevant to our approach is the staggered adoption design where the units are treated at different points of time and remain exposed to the treatment afterward. Several papers have demonstrated that treatment effect heterogeneity across groups and periods can lead to substantially biased estimates when using traditional two-way fixed-effects models. Alternative estimators that are robust to treatment effect heterogeneity have been proposed (see e.g., Callaway and Sant'Anna 2021; Sun and Abraham 2021; Athey and Imbens 2022; Goodman-Bacon, 2021), and the validity of these estimators depends on a parallel trends assumption (Roth and Sant'Anna, 2022). The staggered adoption design takes an "irreversibility of treatment" assumption: $D_{t-1} = 1$ implies $D_t = 1$ (D is an indicator for treatment). Our approach differs in that the treatment indicator in the current period does not depend on its value in the previous period.[1] As such, we allow for multiple treatments on the same units. More importantly, the crossover design replaces the parallel trends assumption with weaker assumptions on homogenous signs and dynamics of the treatment effect. The basic crossover design can be viewed as a specific combination of two DID designs where there is a "reverse comparison" between the two groups. Our approach relies on the idea that in the crossover design any economic trends that might have biased the previous result will have the opposite effect on the subsequent result, and the offset of these two forces will help identify the treatment effect.

The recent development of the DID literature has also pointed out some shortcomings of the traditional pre-trends tests, such as low power and distortions from pre-trends testing (Freyaldenhoven et al., 2019; Roth, 2020; Kahn-Lang and Lang, 2020), and started to address violations of the parallel trends assumption. Three approaches are notable: inferences based on reformulated models that allow for non-parallel pre-treatment trends (Manski and Pepper, 2018; Bilinski and Hatfield, 2019;

[1] For a two-groups, three-periods design, the crossover setting $[(0,0)|(0,1)|(1,0)]$ means that the second group is assigned a treatment in the second period and the first group is assigned a treatment in the third period. This is essentially the same as the staggered adoption setting $[(0,0)|(0,1)|(1,1)]$, where the second group is treated in the second period only but its treated status is noted in the third period (and following periods).

Rambachan and Roth, 2022), use of an observed covariate as the instrument for an unobserved confounder that causes the violation of parallel trends (Freyaldenhoven et al., 2019), and the parallel growth assumption that requires the derivatives of the paths to be parallel (Mora and Reggio, 2019). Similar to these approaches, the crossover design imposes a restriction on possible violations of the parallel trends assumption and focuses on the robustness of inferences to the violations. A key difference of our approach, however, is that we rely on a novel design that can identify spurious estimators without pre-treatment observations. For instance, in the two-groups, three-periods design described above, one can get an estimate of the ATE through the staggered DID approach, but it is difficult to assess the validity of the estimator as the pre-trends cannot be tested. Applying the crossover design, however, one can make an inference on the ATE as well as evaluate the validity of the estimator. The crossover design is particularly useful in empirical settings where the researcher is more confident in homogenous signs of the treatment effect but less confident in the parallel trends. The simplicity and effectiveness of the crossover design make it a useful addition to the existing DID literature.

Second, we contribute to the minimum wages research in China in two ways. To begin with, we present new credible evidence on the wages and employment effects of minimum wages in China. Most previous studies use data before the year 2010 (Jia, 2014; Fang et al., 2021), while in this study we use more recent individual data from CMDS that has several advantages. First, microdata allows us to control for individual-level variables to get efficient estimators. Second, existing studies usually use data that covers mostly urban workers. Our study, however, focuses on rural migrant workers who are more likely to be affected by minimum wages than urban workers. Third, minimum wages in China have increased substantially during the period we analyze (2011-2014), which may allow us to more clearly identify the effects of minimum wages than studies using data from earlier periods.

In addition, our study suggests inappropriately matched comparison groups as an explanation for the mixed findings on the employment effect of minimum wages in China. DID-based Chinese minimum wages studies normally match the treatment and control groups on area-level characteristics (often neighboring areas). However, as many use one-period, cross-sectional data, they are not able to test the parallel trends (e.g., Jia, 2014; Yang and Li, 2016). We follow this approach to choose the treatment and control groups and apply the crossover design. The results show that DID analyses with seemingly comparable groups may possibly come to opposite conclusions in different one-period analyses, which challenges the validity of existing studies that fail to address the parallel trends concern.

The remainder of the paper is organized as follows. Section 2 surveys the minimum wages literature with a focus on the challenge of the parallel trends assumption and discusses the intuition to conduct the crossover design in economic research. Section 3 presents the crossover design and its assumptions and inference procedures. Section 4 briefly describes the setting of minimum wages in China. Section 5 describes the data and model specification. Section 6 presents our main results and examines the robustness of the findings. Section 7 concludes.

2. DID design and minimum wages research: challenge of the parallel trends assumption

The DID design is widely used in the minimum wages literature. The disagreement among studies on the employment effect of minimum wages in the United States is well known (for extensive reviews see Cengiz et al., 2019; Dube, 2019; Neumark and Shirley, 2021). The plausibility of the parallel trends assumption is a critical issue in the debate. For illustration, in a recent debate, Sabia et al. (2012) apply the conventional DID design with CPS-MORG (a subset of CPS) data to examine the effects of an increase of the minimum wage in New York State and find a large, significant, and negative employment effect. In their study, the parallel trends assumption is satisfied (rather than tested) with the use of synthetic controls. However, Hoffman (2016) replicates their study with the full CPS data and finds no evidence of any employment effect. Sabia et al. (2016) reply that Hoffman's conclusions are insufficient as he does not provide valid evidence on the parallel trends before the intervention.

A growing number of studies have also applied the DID design to identify labor market effects of minimum wages in China, the findings of which are mixed.[2] Some scholars find that minimum wages have a significant and negative impact on employment (Ding, 2010; Wang and Gunderson, 2012), whereas other scholars report insignificant employment effects (Mayneris et al., 2018; Wang and Gunderson, 2018; Fang et al., 2021) or significant and positive effects (Ni et al., 2011). It is difficult to explain these mixed findings because none of the studies have implemented a placebo test to support the internal validity of the DID design. The reason is that provinces in China have considerable autonomy and flexibility in setting their minimum wages and they are mandated to adjust minimum wages at least once every two years. Given that large-scale individual-level surveys in China are

[2] In China, Fang and Lin (2015) show that local business cycle variables such as local average GDP growth or unemployment do not predict local minimum wages changes. Therefore, the increase of minimum wages can be taken as exogenous, a condition on which the DID design relies.

normally conducted with at least one-year intervals, once the treatment and control groups are selected to conduct a DID design, it will be difficult to find a pre-treatment period in which neither group adjusted minimum wages to check the parallel trends assumption.

Due to the data limitation for the DID design, some scholars turn to other strategies such as the panel model that directly regresses outcome variables on minimum wages with fixed effects to study minimum wages effects in China. However, these approaches usually use aggregated provincial-level or county-level data with imprecise measures of the minimum wage such as the time-weighted average annual minimum wages, which may introduce measurement errors in the analysis (Wang and Gunderson, 2012; Fang and Lin, 2015). In addition, the consensus has yet to be reached on how to properly control the underlying region-specific time trends, and it is difficult to understand what the counterfactual is in these panel models (Dube et al., 2010; Neumark et al., 2014; Meer and West, 2016; Clemens and Wither, 2019). In the absence of a randomized experiment, alternative research designs should be considered to improve the credibility of the studies.

Crossover design is a traditional experimental design approach in the clinical medicine field (Maclure and Mittleman, 2000) that identifies effects of a treatment by exposing subjects to the treatment separately in different time periods. The most basic crossover design consists of two groups and two periods in which one group is given a treatment in the first period and a dummy treatment (placebo) in the second period whereas the other group is firstly given a placebo and then the treatment in the second period. Compared with the parallel design in which the treatment group and the control group remain unchanged in the second period, the crossover design makes each group serve as its own control which helps reduce the influence of confounding factors and avoid selection bias.

The crossover design in the clinical medicine field, however, cannot be directly applied to economic analysis due to differences in the premises. In clinical trial studies, only two periods are required to conduct a crossover design as the treatment and control groups are already comparable through randomization (the parallel trends assumption is met), and the motivation of the crossover design is to use fewer samples to attain the same level of statistical power or precision as the parallel design. In economic studies, however, researchers cannot design an experiment to achieve the necessary randomness. Our intuition to consider a crossover design is to make robust causal inferences on treatment effects when the parallel trends assumption cannot be tested directly. As a result, three periods are required and the assumptions and inference procedures of the crossover design in our study are substantially different from those in the clinical medicine literature.

Two minimum wages studies relate to the idea of the crossover design. In their famous study, Card and Krueger (1994) take the increase of the minimum wage in New Jersey in April 1992 as a quasi-experiment, in which New Jersey is regarded as the treatment group and Pennsylvania the control group, to examine the employment effect of minimum wages. They find a positive effect of the minimum wage on employment, which, however, has caused a debate due to the lack of the placebo test (Neumark and Wascher, 2000). In response, Card and Krueger (2000) report another case study when the federal minimum wage increased from \$4.25 to \$4.75 during 1996-1997. The increase was binding in Pennsylvania but had no impact on New Jersey since the state's \$5.05 minimum wage already exceeded the new federal standard. As such, Pennsylvania is regarded as the treatment group and New Jersey the control group. The results are similar to those in the previous study and thus are used as support for the validity of their original conclusions. However, the longer time-series results show that the employment growth of Pennsylvania is not proper to infer the counterfactual results of New Jersey (Angrist and Pischke, 2009), and for this reason, the validity of Card and Krueger (1994)'s results remain controversial.

In China's minimum wages research, Ding (2010) notices that Guangdong and Fujian provinces successively increased minimum wages in 2007 and 2008. However, the purpose of the study is not to accurately infer the employment effect of minimum wages but to examine the effect of the implementation of the labor contract law (which took effect in 2008) on the employment effect of minimum wages. The DID estimators in 2007 and 2008 are 0.132 and -0.079,[3] based on which the author concludes that the negative employment effect of minimum wages is exacerbated by the labor contract law. However, the results could be simply due to the violation of the parallel trends assumption. For instance, if the employment growth of the treatment group is higher than that of the control group in 2007, then even if the treatment effect did not exist, the DID estimators would show opposite signs when the treatment group and the control group were reversed in 2008. Therefore, the credibility of the conclusions is in question.

In short, to our knowledge, the application of the crossover design in economic studies has not been well discussed. This is a key purpose of our study.

[3] The author uses seasonal firm-level data and chooses the seasons before and after the minimum wages hikes.

3. Crossover design

3.1. Setup

A typical crossover design has three time periods $T \in \{0,1,2\})$ and two groups $G \in \{A, B\}$. In period 0, neither group is treated. One group is assigned the treatment in period 1 and the other group is assigned the treatment in period 2. The binary treatment variable D_P indicates whether a group is treated in period P.

Based on the potential outcomes framework (Rubin, 1974; Robins, 1986), let $Y_{GT}(D_1, D_2)$ (as all units have $D_0 = 0$ in period 0, we omit this from the notation) denote the potential outcomes for group G in period T if units were to follow the treatment path (D_1, D_2).

Assumption 1. The potential outcomes for group G at a given time T are not affected by future assignments: $Y_{G0} = Y_{G0}(D_1, D_2)$, $Y_{G1}(D_1, .) = Y_{G1}(D_1, D_2)$.

Assumption 1—usually referred to as the no-anticipation assumption— is commonly used in the DID literature (Callaway and Sant'Anna 2021). Under assumption 1, the potential outcomes can be defined as:

$$Y_{GT}(D_1, D_2) = \left\{ \begin{array}{lll} Y_{G0} & & T = 0 \\ Y_{G1}(D_1) = \begin{cases} Y_{G1}(1) & if\ D_1 = 1 \\ Y_{G1}(0) & if\ D_1 = 0 \end{cases} & & T = 1 \\ Y_{G2}(D_1, D_2) = \begin{cases} Y_{G2}(1,1) & if\ D_1 = D_2 = 1 \\ Y_{G2}(1,0) & if\ D_1 = 1, D_2 = 0 \\ Y_{G2}(0,1) & if\ D_1 = 0, D_2 = 1 \\ Y_{G2}(0,0) & if\ D_1 = D_2 = 0 \end{cases} & & T = 2 \end{array} \right\}. \tag{1}$$

In period 0, there is only one outcome for each group as no treatment occurs. In period 1, there are two potential outcomes for each group: $Y_{G1}(1)$ if treated and $Y_{G1}(0)$ if untreated in period 1. In period 2, there are four potential outcomes for each group: $Y_{G2}(1,1)$ if treated in periods 1 and 2, $Y_{G2}(1,0)$ if treated in period 1 and untreated in period 2, $Y_{G2}(0,1)$ if untreated in period 1 and treated in period 2, and $Y_{G3}(0,0)$ if untreated in both periods.

Suppose that A is treated in period 1 and B is treated in period 2. Then we observe $Y_{GT}(1,0)$ for group A and $Y_{GT}(0,1)$ for group B. Specifically, Y_{A0}, $Y_{A1}(1)$, $Y_{A2}(1,0)$ is observed for group A and Y_{B0}, $Y_{B1}(0)$, $Y_{B2}(0,1)$ is observed for group B. Our parameters of interest are the ATE on group A: $E[Y_{G1}(1) - Y_{G1}(0)|G = A]$ and group B: $E[Y_{G2}(0,1) - Y_{G2}(0,0)|G = B]$.

Given the notation above, we can define the first DID estimator by comparing the evolution of the mean outcome in the two groups between periods 0 and 1:

$$\begin{aligned}\delta_1 &= E[Y_{G1}(1) - Y_{G0}|G = A] - E[Y_{G1}(0) - Y_{G0}|G = B] \\ &= E[Y_{G1}(1) - Y_{G1}(0)|G = A] + E[Y_{G1}(0) - Y_{G0}|G = A] - E[Y_{G1}(0) - Y_{G0}|G = B] \\ &= \tau_{At} + \nu_1,\end{aligned} \tag{2}$$

where $\tau_{A1} = E[Y_{G1}(1) - Y_{G1}(0)|G = A]$ denotes the average treatment effect of the first treatment on group A in period 1, and $\nu_1 = E[Y_{G1}(0) - Y_{G0}|G = A] - E[Y_{G1}(0) - Y_{G0}|G = B]$ denotes the difference between the treatment and control groups in the time trends of the outcome variable in the absence of treatment for periods 0 and 1. It shows that the first DID estimator equals to the ATT plus the time trends difference. Here note that under the parallel trends assumption: $\nu_1 = 0$.[4]

Similarly, the second DID estimator for periods 1 and 2 is:

$$\begin{aligned}\delta_2 &= E[Y_{G2}(0,1) - Y_{G1}(0)|G = B] - E[Y_{G2}(1,0) - Y_{G1}(1)|G = A] \\ &= E[Y_{G2}(0,1) - Y_{G2}(0,0)|G = B] + E[Y_{G2}(0,0) - Y_{G1}(0)|G = B] \\ &\quad - E[Y_{G2}(1,0) - Y_{G1}(1)|G = A].\end{aligned}$$

Note that:

$$\begin{aligned}E[Y_{G2}(1,0) - Y_{G1}(1)|G = A] &= E[Y_{G2}(1,0) - Y_{G2}(0,0)|G = A] \\ &\quad - E[Y_{G1}(1) - Y_{G1}(0)|G = A] + E[Y_{G2}(0,0) - Y_{G1}(0)|G = A].\end{aligned}$$

Let $\tau_{B2} = E[Y_{G2}(0,1) - Y_{G2}(0,0)|G = B]$ denote the average treatment effect of the second treatment on group B in period 2, $\tau_{A2} = E[Y_{G2}(1,0) - Y_{G2}(0,0)|G = A]$ denote the average treatment effect of the first treatment on group A by time period 2, and $\nu_2 = E[Y_{G2}(0,0) - Y_{G1}(0)|G = B] - E[Y_{G2}(0,0) - Y_{G1}(0)|G = A]$ denote the time trends difference between the treatment and control groups for periods 1 and 2. We have:

$$\delta_2 = \tau_{B2} + \tau_{A1} - \tau_{A2} + \nu_2. \tag{3}$$

Note that the second DID estimator has a different form than the first one as it may be impacted by the treatment on group A in period 1.

3.2. *Identifying assumptions and inference*

In the following, we first consider assumptions under which we can make an inference on the direction of the treatment effect, and then assumptions for the inference on the magnitude of ATE. Further, we discuss the performance of the crossover design when the treatment effect changes over

[4] In the expression of ν, $E[Y_{Gt}(0)|G = A]$ is not observable. Therefore, it is not possible to directly test the parallel trends assumption. Instead, one can check the pre-treatment trends of the outcome variable in the two groups to support the validity of this assumption. The pre-trend test itself is based on the extrapolation assumption that if the outcome variable has parallel trends in the two groups before the treatment, the parallel trends will persist during the treatment. Rambachan and Roth (2022) refer to it as one type of smoothness restrictions. The extrapolation assumption usually holds when the treatment is exogenous.

time.

Assumption 2. In the absence of treatment, the sign of the time trends difference between the two groups does not change over time: $\nu_1\nu_2 < 0 \veebar \nu_1 = \nu_2 = 0$**.**

Assumption 2 requires either the means of group A and group B have parallel trends ($\nu_1 = \nu_2 = 0$) or their relative growth rank do not change. It allows for the possibility that the time trends of the outcomes in the two groups may be different, though with restrictions. For instance, in the absence of treatment, if the outcome variable of group A grows faster than that of group B between periods 0 and 1, it should also grow faster between periods 1 and 2. This assumption is plausible when the stochastic shocks are small relative to the differences in the time trends. In the minimum wages context, Assumption 2 is likely to hold when the two groups (areas) are matched with similar economic conditions and the time interval between periods is not long (because the impact of economic cycles is less likely to be systematically different in the shorter run). This is a key assumption of the crossover design as it relaxes the parallel trends assumption.

Assumption 3. Homogeneous sign of treatment effect between groups in the period that the treatment occurs: $\tau_{A1} = \tau_{B2} = 0 \veebar \tau_{A1}\tau_{B2} > 0$.

Assumption 3 implies that the treatment does not have an impact on either group ($\tau_{A1} = \tau_{B2} = 0$) or impacts the two groups in the same direction ($\tau_{A1}\tau_{B2} > 0$). The intuition is that since the two groups are comparable, the treatment effect should be consistent in the sign. This assumption relaxes the homogenous treatment effect assumption as it allows for a specific version of heterogeneous treatment effects.

Assumption 4a. Stable treatment effect over time: τ_G does not vary with T.

Assumption 4b.: Stable treatment effect over time when the treatment effect is zero in the period that the treatment occurs, and bounded dynamic treatment effect when the treatment effect grows over time: $\tau_{A2} = 0$ if $\tau_{A1} = 0 \wedge |\tau_{B2}| \geq |\tau_{A1} - \tau_{A2}|$ if $\tau_{A1}(\tau_{A1} - \tau_{A2}) < 0$.

Assumption 4a implies: $\tau_{A1} = \tau_{A2}$, then we have: $\delta_2 = \tau_{B2} + \nu_2$, that is, the DID estimator for periods 1 and 2 will not be influenced by the treatment given in period 1.[5] Assumption 4a, commonly

[5] Assumption 4a is equivalent to the "no carryover effect assumption" in the crossover design of clinical trials. A carryover effect is defined as the effect of the treatment from the previous period on the response in the current period. The purpose of both assumptions is to ensure that the second treatment effect is not influenced by the first one. The difference is that in clinical trials it is assumed that the treatment has an abrupt onset that completely resolves by the next period, that is, there is a washout period between the two trials. In our crossover design, however, we assume that the treatment effect is stable over time, which is more likely to hold when the trials are continuous in time.

required in many extensions of the DID design (De Chaisemartin and D' Haultfoeuille, 2018; Athey and Imbens, 2022), is necessary to obtain our ATE estimator. From the practical point of view, Assumption 4b is a much weaker version of Assumption 4a. When the treatment effect is zero in the period that the treatment occurs, Assumption 4b suggests that the treatment has no impact over time, which excludes the situation where the treatment effect needs long time to emerge. When the treatment has an instant non-zero impact and the impact increases over time ($\tau_{A1}(\tau_{A1} - \tau_{A2}) < 0$), Assumption 4b requires that the treatment effect do not change too rapidly such that the sign of the second treatment's effect will not be reversed by the residual effect of the first treatment.

When Assumptions 1, 2, 3, and 4 (4a or 4b) are satisfied, the inference rules for the direction of treatment effects are given as follows:

Proposition 1. *If the two DID estimators have the same sign and neither is zero, then the treatment effect is non-zero and the sign of the treatment effect in the period that the treatment occurs is the same as the sign of the DID estimators:*

If $\delta_1\delta_2 > 0$, *then* $\tau_G \neq 0$ *and* $\tau_G\delta_{\mathrm{p}} > 0$.

Suppose $\delta_1\delta_2 > 0$. Under Assumption 3 and Assumption 4b, if the treatment effect is zero $\tau_G = 0$, we have $\tau_{A1} = \tau_{A2} = \tau_{B2} = 0$. Then $\delta_1\delta_2 = (\tau_{A1} + \nu_1)(\tau_{B2} + \tau_{A1} - \tau_{A2} + \nu_2) = \nu_1\nu_2 > 0$, which violates Assumption 2. Therefore, $\tau_G \neq 0$.

As $\tau_G \neq 0$, under Assumption 3, $\tau_{A1}\tau_{B2} > 0$. When the treatment has an instant non-zero impact and the impact decreases over time, we have $\tau_{A1}(\tau_{A1} - \tau_{A2}) > 0$, and then $\tau_{A1}(\tau_{B2} + \tau_{A1} - \tau_{A2}) = \tau_{A1}\tau_{B2} + \tau_{A1}(\tau_{A1} - \tau_{A2}) > 0$. When the treatment has an instant non-zero impact and the impact increases over time, that is, $\tau_{A1}(\tau_{A1} - \tau_{A2}) < 0$, under Assumption 4b, we have $\tau_{A1}(\tau_{B2} + \tau_{A1} - \tau_{A2}) \geq 0$. When the treatment has an instant non-zero impact and the impact is stable over time, then $\tau_{A1} = \tau_{A2}$ and $\tau_{A1}(\tau_{B2} + \tau_{A1} - \tau_{A2}) > 0$. Therefore, $\tau_{A1}(\tau_{B2} + \tau_{A1} - \tau_{A2}) \geq 0$, and then $\tau_{A1}(\tau_{B2} + \tau_{A1} - \tau_{A2}) + \delta_1\delta_2 = \tau_{A1}\delta_2 + (\tau_{B2} + \tau_{A1} - \tau_{A2})\delta_1 + \nu_1\nu_2 > 0$. Under Assumption 2, $\nu_1\nu_2 \leq 0$, then $\tau_{A1}\delta_2 + (\tau_{B2} + \tau_{A1} - \tau_{A2})\delta_1 > 0$. As $\tau_{A1}\delta_2 \cdot (\tau_{B2} + \tau_{A1} - \tau_{A2} + \nu_2)\delta_1 > 0$, we have $\tau_{A1}\delta_2 > 0$. Then τ_{A1} and τ_{B2} should have the same sign as δ_1 or δ_2, that is, $\tau_G\delta_{\mathrm{p}} > 0$. The intuition is that the bias generated by the difference in time trends between the two groups is small relative to the treatment effect.

Proposition 2. *If both of the two DID estimators are zero, then the parallel trends assumption holds and the treatment has no impact on either group:*

If $\delta_1 = \delta_2 = 0$, *then* $\nu_1 = \nu_2 = 0$ *and* $\tau_A = \tau_B = 0$.

Under Assumption 4a, when $\delta_1 = \tau_A + \nu_1 = 0$, $\delta_2 = \tau_B + \nu_2 = 0$, we have: $\tau_A = -\nu_1$, $\tau_B = -\nu_2$. Under Assumption 2, $\tau_A\tau_B = \nu_1\nu_2 \le 0$. Under Assumption 3, $\tau_A\tau_B \ge 0$. Therefore, $\tau_A\tau_B = \nu_1\nu_2 = 0$. Under Assumptions 2 and 3, we have: $\nu_1 = \nu_2 = 0$ and $\tau_A = \tau_B = 0$.

Proposition 3. *If one DID estimator is zero while the other is not, then the parallel trends assumption does not hold, but the treatment effect is non-zero and has the same sign as the non-zero DID estimator:*

If $\delta_1 = 0 \veebar \delta_2 = 0$ *and* $\delta_p \neq 0$, *then* $\nu_1\nu_2 \neq 0$, $\tau_G \neq 0$ *and* $\tau_G\delta_{\mathrm{p}} > 0$.

Suppose $\delta_1 = 0$, $\delta_2 > 0$. Under Assumption 4a, we have $\tau_A = -\nu_1$, $\tau_B > -\nu_2$. Under Assumption 2, if $\nu_1 = 0$, then $\nu_2 = 0$, $\tau_A = 0$, $\tau_B > 0$, which violates Assumption 3; if $\nu_1 > 0$, then $\tau_A\tau_B < \nu_1\nu_2 < 0$, which also violates Assumption 3. Therefore, we have $\nu_1 < 0$, $\tau_A > 0$. Further, under Assumptions 2 and 3, $\nu_1\nu_2 < 0$, $\tau_B > 0$. Putting together, we have $\nu_1\nu_2 \neq 0$, $\tau_G \neq 0$ *and* $\tau_G\delta_2 > 0$.

Proposition 4. *If the two DID estimators have different signs and neither is zero, then the parallel trends assumption does not hold, and it is not possible to make an inference on the treatment effect:*

If $\delta_1\delta_2 < 0$, *then* $\nu_1\nu_2 \neq 0$, $\tau_G =?$

When $\delta_1\delta_2 < 0$, the treatment effect may or may not exist. If the treatment effect does not exist $\tau_G = 0$, under Assumption 4a, $\delta_1 = \tau_A + \nu_1 = \nu_1$, $\delta_2 = \tau_B + \nu_2 = \nu_2$. Then $\delta_1\delta_2 = \nu_1\nu_2 < 0$, which violates the parallel trends assumption. If the treatment effect is non-zero $\tau_G \neq 0$, under Assumption 4a, we have $\delta_1\delta_2 = (\tau_A + \nu_1)(\tau_B + \nu_2) = \tau_A\tau_B + \tau_A\nu_2 + \tau_B\nu_1 + \nu_1\nu_2 < 0$. If $\nu_1\nu_2 = 0$, then under Assumption 2, $\nu_1 = \nu_2 = 0$. We have $\delta_1\delta_2 = \tau_A\tau_B < 0$, which violates Assumption 3. Therefore, $\nu_1\nu_2 \neq 0$, that is, the parallel trends assumption does not hold. In either case we are unable to make an inference on the treatment effect.

By far we have tried to make an inference only on the direction of the treatment effect, but one can answer more questions at the cost of more assumptions. To get the ATE, we introduce two more assumptions.

Assumption 5. In the absence of treatment, the growth of the time trends difference between the two groups does not change over time: let Δt be the sampling time interval, then $\nu(\Delta t) = \bar{C}\Delta t$ where $\bar{C}$ is a constant.

Assumption 6. Homogeneous treatment effect between groups in the period that the treatment occurs: $\tau_{A1} = \tau_{B2} = \tau$.

Under Assumptions 4a, 5, and 6, we have:

Proposition 5. *If the two sampling time intervals in the crossover design are equal, then the ATE of the treatment is:* $\hat{\hat{\tau}} = \frac{\delta_1+\delta_2}{2}$.

Under Assumption 4a, the definition of the ATE is: $\bar{\tau} = \tau_A \cdot P(G = A) + \tau_B \cdot P(G = B)$. Under Assumptions 6 we have: $\bar{\tau} = \tau\big(P(G = A) + P(G = B)\big) = \tau$, $\delta_1 = \tau + \nu_1$, $\delta_2 = \tau + \nu_2$. If the two sampling time intervals are equal, under Assumption 5 $\nu_1 = -\nu_2 = \nu_{\Delta t}$. Then we have $\delta_1 - \nu_{\Delta t} = \delta_2 + \nu_{\Delta t} = \tau$, where $\nu_{\Delta t}$ is not observable. It follows that:

$$\hat{\hat{\tau}} = \frac{\delta_1 + \delta_2}{2} = \frac{\tau + \nu_{\Delta t} + \tau - \nu_{\Delta t}}{2} = \tau = \bar{\tau}.$$

Therefore $\hat{\hat{\tau}}$ is an unbiased estimator for $\bar{\tau}$. The standard error for the ATE estimator is: $S_{\hat{\hat{\tau}}} = \frac{\sqrt{S_{\delta_1}{}^2+S_{\delta_2}{}^2}}{2}$.

3.3. Discussion

In most of the analysis above, we assume that the treatment has a persistent effect. However, there are situations where Assumption 4a is implausible, because the treatment effect may have dynamic features and vary over time. Consequently, the bias of DID estimators will be sensitive to the relative duration of time (Meer and West, 2016). Therefore, it is necessary to further examine the application of the crossover design when Assumption 4a does not hold.

We will discuss two specific cases in the following. Let t_0, t_1, t_2 be the first, second, and third period in which the outcome variable is observed. For simplicity, suppose that group A and group B have parallel trends with regard to the outcome variable ($\nu_1 = \nu_2 = 0$) and the treatment effect is homogeneous ($\tau_A = \tau_B = \tau$). Let C denotes the counterfactual group; then the DID estimator equals the difference between the outcome variable of the treatment group and its counterfactual on the observation date.

Case 1. The treatment has both an instant and a dynamic impact on the outcome variable. Assume that the treatment effect is positive and decreases over time ($\tau_{A1} > \tau_{A2}$). As shown in Figure 1(a), when the sampling dates t_1 and t_2 are both right after the treatments, the first DID estimator captures the treatment effect ($\delta_1 = \tau$), but the second DID estimator is upward biased relative to the true effect ($\delta_2 > \tau$). Figure 1(b) shows that when the observation dates t_1 and t_2 both succeed the two treatments, we have $\delta_1 < \tau < \delta_2$. Consequently, the relative size of the estimators does not provide information on whether the treatment effect has increased or decreased.

Case 2: The treatment only impacts the growth rate of the outcome variable. Assume that the treatment has a negative impact on the growth of the outcome variable ($\tau_{At} > \tau_{A(t+1)}$). Figure 1(c) shows that when the sampling dates t_1 and t_2 are both right after the treatments, the two DID estimators are zero ($\delta_1 = 0$) and positive ($\delta_2 > 0$), respectively. Figure 1(d) indicates that when the sampling dates t_1 and t_2 both succeed the two treatments, the two DID estimators are negative ($\delta_1 < 0$) and positive ($\delta_2 > 0$), respectively. The results suggest that even if the parallel trends assumption holds for the two groups, the DID estimators may still produce conflicting results: $\delta_1\delta_2 \leq 0$.

<< ***Insert Figure 1 here*** >>

The two cases illustrate that when Assumption 4a does not hold, the inference becomes complicated and both Proposition 3 and Proposition 4 fail. As a result, we cannot make an inference on the parallel trends or the treatment effect when the two DID estimators are different in sign.

In addition, there are situations where we may suspect potential violations of other assumptions of our design. First, we may suspect that the difference in trends between the two groups reverses over time such that Assumption 2 is violated: e.g., Ashenfelter's dip predicts a negative difference in trends at first and a positive difference afterward. Second, it may be that different groups have heterogeneous treatment effects of opposite signs such that Assumption 3 does not hold. The estimators proposed by Callaway and Sant'Anna (2021) and De Chaisemartin and D' Haultfoeuille (2020) can allow for arbitrary treatment effect heterogeneity, but they both require the parallel trends assumption. There is a trade-off between the homogeneity assumption and the parallel trends assumption.

Compared to extant literature, the crossover design is particularly useful in empirical settings where the researcher is more confident in homogeneous signs of the treatment effect but less confident in the parallel trends, especially when (i) there are no groups that remain untreated during two consecutive periods such that the pre-trends cannot be tested directly; (ii) the pre-trends can be tested, but the pre-trends periods are not long enough to provide robust support for the parallel trends assumption; (iii) there are multiple treatments on the same unit.

To conclude, the analysis above highlights that in the crossover design the violation of the parallel trends assumption does not necessarily fail the identification of the treatment effect, and the validity of inference does not necessarily rely on the pre-trend test. It shows that: (1) if Assumptions 1, 2, 3, and 4b are satisfied, then Proposition 1 holds; (2) if Assumptions 1, 2, 3, and 4a are satisfied, then Propositions 2, 3, and 4 hold; (3) if Assumptions 1, 4a, 5, and 6 are satisfied, then Proposition 5 holds. Note that Proposition 1 is the most robust one as it needs the least restrictive assumptions.

The crossover design above can be extended to accommodate cases for more than three periods with multiple treatments on the same unit, as long as we have a data structure like $[(0,0)|(0,1)|(1,0)|(0,1)|\ldots]$, and the ATE should be calculated by the average of two adjacent estimators $\hat{\bar{\tau}} = \left(\delta_{p-1} + \delta_p\right)/2$. However, as the number of periods increases, the utility of crossover design to practical analyses becomes lower because for non-experimental data the assumptions are less likely to hold in longer periods. This is a limitation for all DID extensions that involve multiple periods.

4. Minimum wages in China

This section briefly describes the institutional background of the minimum wage in China.[6] China's first minimum wages regulation, i.e., the Enterprise Minimum Wages Regulations, was issued by the Ministry of Labor in 1993, which establishes the standard of minimum wages on a monthly rather than hourly basis. Due to the huge difference in living standards across regions, there is no universal minimum wage levels for the entire nation. Instead, provinces have considerable autonomy and flexibility in setting their minimum wages. In most provinces, cities and counties are sorted into several tiers (from two to five tiers) according to local economic conditions. The government then consults the trade union federation and the employer representative to decide on the level of minimum wages for each tier. In the early years of its implementation, local governments had little incentive to raise or enforce minimum wages due to their top priority of economic development. As a result, the level of minimum wages was extremely low and the enforcement was very weak till early 2000s (Liu, 2009).

To strengthen the minimum wages regulation, the Ministry of Labor and Social Security issued the Provisions on Minimum Wages in 2004, which has the following important features: first, local governments are mandated to adjust minimum wages at least once every two years; second, in addition to monthly minimum wages for full-time workers, hourly minimum wages are introduced to cover part-

[6] See e.g., Ye et al. (2015) and Fang and Lin (2015) for more details about minimum wages regulations in China.

time workers;[7] third, the coverage of minimum wages is extended to all employees; fourth, the penalties for violations of minimum wages are increased substantially from 20%~100% of wages owned to employees to 100%~500%.

Since the implementation of the 2004 regulation, both the frequency and magnitude of changes in minimum wages have been substantial. Although the adjustment of minimum wages was suspended in 2009 due to the global financial crisis, a new wave of minimum wage hikes followed soon after the recession. Among 32 regions in mainland China (31 provincial-level administrative regions and Shenzhen city), 30, 25, 25, 27, 19, and 27 regions adjusted minimum wages each year between 2010 and 2015, and the average increase was 23%, 22%, 20%,18%,14%, and 15%, respectively.

In 2016, the Ministry of Labor and Social Security lowered the frequency of minimum wages adjustment from every two years to every two to three years due to the downward pressure on China's economy. Consequently, the frequency and magnitude of minimum wages adjustment decreased after 2015. For instance, only 9 provinces adjusted minimum wages in 2016 and the average increase was 11%. The conservative policy change reflects the government's concern of the negative effect of minimum wages on the economy. However, existing findings regarding the effects of minimum wages in China are mixed. Moreover, the past decade has witnessed China's entering a qualitatively different stage of economic development and moving from unlimited labor supply to a new stage of labor shortage, which may possibly impact the effects of minimum wages. The China's minimum wages research using post-2010 data, however, has been very rare. Our study therefore aims to provide robust and updated evidence on minimum wages effects during 2010-2015, a period of rapid and substantial minimum wages hikes in China.

5. Data and model specification

5.1. *Data*

Our data are drawn from the individual repeated cross-sectional data of CMDS, which is conducted annually by the Chinese National Health Commission using the stratified multi-stage probability proportionate to size (PPS) method for sampling. Launched in 2009, the CMDS has become the most nationally representative survey of rural migrant workers. The surveys consist of 128, 159, 196, and

[7] Since employers are not required to provide non-wage benefits for part-time workers, the hourly minimum wage is set higher than hourly-adjusted monthly minimum wages (monthly minimum wages divided by 174 regular working hours per month).

200 thousand individuals in 2011, 2012, 2013, and 2014, respectively.[8] The respondents are floating migrants aged 15-59 who do not have local household registration (*Hukou*) and live in their current place for more than one month. The CMDS was conducted in July in 2011 and in May each year between 2012 and 2014. An advantage of the CMDS is that its questions about wages and employment are accurate to the specific month rather than the specific year, which allows us to pinpoint whether the respondents are affected by the latest minimum wages increase on the dates of the survey. Specifically, a worker is defined as potentially affected (treatment group) if there is a minimum wages hike in the worker's province between the previous and current survey. It should be noted that it is inappropriate to define minimum wages workers as those whose wages are just around the minimum wages (which is common in the hourly minimum wages literature) in a context where minimum wages are on a monthly basis. The monthly minimum wages in China are based on 40 hours of work per week, but the average weekly working hours of migrant workers are normally much longer than 40. Moreover, overtime wages vary: 1.5 times of basic wages on weekdays, 2 times on weekends, and 3 times on holidays. The spillover effect of minimum wages on workers with wages higher than minimum wages makes the identification of minimum wages workers even more complicated or even impossible. Another advantage of the CMDS is that the sampling time intervals are largely the same in these years, which is very desirable for the crossover design.

Information on minimum wages used in this study comes from the Ministry of Human Resources and Social Security and the websites of the local governments. Table 1 presents the status of minimum wages adjustment in each province during 2011-2014. In any of these years, status of adjustment equals 1 if a province adjusted the minimum wage between the previous and current CMDS surveys, and 0 otherwise.

<< *Insert Table 1 here* >>

We therefore match the provinces first on three criteria: (1) the same economic region (the China National Bureau of Statistics divides China into four major economic regions: eastern region, central region, western region, and northeastern region); (2) similar growth rates of GDP per capita; (3) similar

[8] For simplicity, we only use the 2011-2014 CMDS data to illustrate the application of the crossover design. We repeat our analysis using data from the 2017 CMDS survey in the robustness check.

levels of minimum wages, i.e., the minimum wage standard of the control group is within the range of the pre- and post-treatment minimum wages of the treatment group. Further, given that the crossover design requires the two provinces be treated alternately and that shorter intervals are more appropriate because the impact of economic cycles is more likely to be systematically different in the longer run, our matching ends with three groups of provinces: Jiangxi and Anhui in the central region, Sichuan and Chongqing in the western region, and Jilin, Liaoning, and Heilongjiang in the northeast region (as shown in Figure 2).

<< ***Insert Figure 2 here*** >>

Table 2 presents GDP per capita and minimum wages standards of the matched provinces between 2011 and 2014. It shows that the minimum wages standards in each group of provinces are close with each other in these years. Moreover, Anhui and Jiangxi in the central region had very similar levels and growth rates of GDP per capita and they alternately adjusted minimum wages in 2013 and 2014;[9] in the western region, Sichuan (province) and Chongqing (city) alternately increased minimum wages between 2011 and 2013, and Chengdu—the capital city of Sichuan—is selected as the comparison group of Chongqing for its higher comparability than the entire Sichuan province; in the northeast region Jilin and Heilongjiang (J&H hereafter) are combined as the comparison group of Liaoning because J&H and Liaoning alternately increased minimum wages between 2011 and 2013 and these provinces had very similar levels and growth rates of GDP per capita. The provinces in the eastern region are not included in our analysis because there are no provinces groups that meet our criteria.[10]

<< ***Insert Table 2 here*** >>

[9] Although the minimum wage standard in Jiangxi province increased by 41% from 2012 to 2013, this was mainly due to the change of statistical criteria. Specifically, the minimum wage standard of Jiangxi province in 2012 did not include social insurances contributions paid by individuals, which, however, were added in 2013. Adjusting for this difference, the actual increase of the minimum wage in these two years was about 13.6%. After 2013, Jiangxi and Anhui have the same statistical criteria for minimum wages.

[10] Fujian and Guangdong provinces in the eastern region are often taken as the comparison groups in DID-based Chinese minimum wages studies (e.g., Ding, 2010). As shown in Table 2, Guangdong province did not adjust minimum wages during 2011-2013 while Fujian province increased the minimum wage from 2012 to 2013. Based on the previous discussion, we can implement the DID design with a placebo test. We take Fujian province as the treatment group and Guangdong province (except Shenzhen City) as the control group and use the 2011-2012 CMDS data to conduct the placebo test. However, the results do not support the parallel trends assumption, which suggests that the two provinces are not comparable groups in our study.

In China, rural migrant workers (according to the CMDS data, 84.77% of the migrants are rural migrant workers) are the main group affected by the minimum wage due to their relatively low level of human capital. We limit our sample to rural migrant workers aged 16-59. We exclude individuals whose employment status are employers or self-employed based on answers to the question “What is your current employment status”, keeping those who are employees or unemployed. We also exclude those who are students, retired or unable to work based on answers to the question “What are the main reasons for not working”. Employment is defined as 1 if one is an employee with reported non-zero wages and 0 otherwise. For those who are employees, wages are measured as monthly wages according to answers to the question “What is your wages in last month (or last employment)”. Outliers of monthly wages below the bottom one percentile and above the upper 99 percentile are winsorized.

Theoretically, the effect of minimum wages is smaller for high-skilled than low-skilled workers. Following Neumark and Wascher (2008), we divide our sample into low-skilled (junior high school and below) and high-skilled (senior high school and above) groups according to workers’ education levels. Because an appropriate specification should at least ensure that the wage effect of minimum wages in the high-skilled group is smaller than that in the low-skilled group (Cengiz et al., 2019), regressions by skill groups can provide additional falsification tests.

Table 3 presents summary statistics of the characteristics of rural migrant workers by provinces and skill groups. Results of t-tests show that mean differences of monthly wages, employment, and most other characteristics of workers between the matched provinces are statistically different at the 1% level. As human capital and other individual characteristics are important factors that influence rural migrant workers' wages and employment, it is necessary to control these variables in regressions to better identify the effects of minimum wages.

<< ***Insert Table 3 here*** >>

5.2. *Model specification*

As noted earlier, the basic crossover design can be viewed as a specific combination of two DID designs. According to Angrist and Pischke (2009), the DID wage equation is set as follows:

$$Y_{icdt} = a + \gamma POST_t + \delta(TREAT_d \times POST_t) + \theta X_{idt} + City_c + \varepsilon_{icdt}. \quad (4)$$

Where the subscript i, c, d, t denotes individuals, cities, groups, and time, respectively. The dependent variable Y_{icdt} is the log of monthly wages. $TREAT_d$ is a dummy variable which equals one for the group that increases the minimum wage and $POST_t$ is a dummy variable with one indicating the date after the increase of minimum wages. δ denotes the DID estimator that captures the ATT of minimum wages. According to Meyer (1995), one can improve the validity of the estimator and relax the parallel trends assumption to a certain extent by controlling for exogenous observable variables that affect the dependent variable. Therefore, we control for years of education, years of post-education experience and its square, gender (1 denotes male and 0 denotes female), marriage (1 denotes married and 0 denotes others), the interaction of marriage and gender X_{idt} at the individual level, and city (prefectural level) fixed effects $City_c$.

The employment equation is set as follows:

$$\Pr(z_{icdt} = 1) = \Phi(a + \gamma POST_t + \delta(TREAT_d \times POST_t) + \theta X_{idt} + City_c) \tag{5}$$

Where the dependent variable is a dummy variable that equals one if the individual is employed, $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The control variables are the same as those in Eq. (4) except that the variable post-education experience is replaced by age. In the nonlinear model, the sign of δ indicates the sign of the treatment effect, but the size itself has no economic meaning. Therefore, we further calculate the treatment effect following the method proposed by Puhani (2012):

$$\begin{aligned} \tau(TREAT_d = 1, POST_t = 1, X) = {} & \Phi(a + \gamma + \delta + \theta X_{idt} + City_c) \\ & -\Phi(a + \gamma + \theta X_{idt} + City_c). \end{aligned} \tag{6}$$

6. Regression results

6.1. *Impact of minimum wages on wages*

The significant wages effect of minimum wages is a necessary but insufficient condition for the existence of employment effect (Manning, 2013). In the following analysis, we first employ the crossover design to examine the wages effect of minimum wages, with specific attention to whether the wages effect exists and whether the treatment and control groups are comparable. We then analyze the employment effect of minimum wages.

Table 4 presents the results of estimation of Eq. (4) for high-skilled rural migrant workers. The regression coefficients of the individual-level control variables are all significant, in line with the expectations of economic theories. The coefficients of $POST_t$ are all positive and significant, indicating that the rural migrant workers' wages have been rising over time.

<< *Insert Table 4 here* >>

The coefficients of the interaction terms are the parameters of interest which are all insignificant in the central and western regions. According to Proposition 2, we can deduce that (1) the parallel trends assumption holds and (2) the increase of minimum wages has no impact on wages of high-skilled rural migrant workers.

However, in the northeastern region the two coefficients of the interaction terms are both significant but different in sign. According to Proposition 4, we can deduce that the parallel trends assumption is violated. If the conclusion from the central and western regions, namely wages of high-skilled rural migrant workers are not affected by the increase of minimum wages, is generalizable, then the results from the northeastern region may be due to that the growth rate of wages of high-skilled migrant workers in Liaoning is lower than that of J&H (Liaoning is the treatment group in 2011-2012 and the control group in 2012-2013). According to Proposition 5, we can still make an inference on the ATE through the aggregation of the two coefficients and expect $\delta_A + \delta_B = 0$. By calculation, $\delta_A + \delta_B = -0.024$ ($se = 0.049$, insignificant at 10% level), which is in line with the expectation. In other words, the significant coefficients are mainly due to the time trends difference between the provinces rather than the treatment effect of minimum wages.

The analysis above shows that for high-skilled rural migrant workers the parallel trends assumption of wages is met in the central and western regions but not in the northeastern region. Whether the findings apply to low-skilled rural migrant workers needs further verification. Table 5 presents the results for low-skilled rural migrant workers. In the central and western regions, the coefficients of the interaction terms are all positive and significant, which, according to Proposition 1, indicates that the increase of minimum wages has a significant and positive impact on wages of low-skilled rural migrant workers. Therefore, in these regions the minimum wage has a greater impact on wages of low-skilled rural migrant workers than that of high-skilled rural migrant workers, suggesting that our model specification passes the falsification test implied by Cengiz et al. (2019). Further, according to Proposition 5, it can be estimated that raising minimum wages increases average wages of low-skilled rural migrant workers in the central and western regions by 4.55% and 6.45%, respectively.

<< Insert Table 5 here >>

In the northeastern region, it is very likely that wages of low-skilled rural migrant workers also violate the parallel trends assumption, which is confirmed by the results. As Table 5 shows, one DID estimator is positive and significant while the other is close to zero and insignificant, suggesting that researchers who use the 2011-2012 data or 2012-2013 data alone would come to different conclusions. According to Proposition 3, we can deduce that the treatment effect of minimum wages on wages is positive and significant. According to Proposition 5, it can be estimated that the ATE of minimum wages on wages is $\bar{\tau} = \frac{\delta_A + \delta_B}{2} = 0.025$ ($se = 0.013$, significant at 10% level), which means, the increase of minimum wages also has a positive impact on wages of low-skilled workers in the northeastern region. We also operate Jilin and Heilongjiang as the comparison group of Liaoning separately and the results are similar. These results indicate that Jilin and Heilongjiang are not proper comparison groups for Liaoning for the purpose of this study.

6.2. *Impact of minimum wages on employment*

The results of the wage equation show that the matched provinces in the central and western regions can well meet the premise of the crossover design. Therefore, the analysis of the employment effect of minimum wages mainly focuses on these two regions and the low-skilled group. Table 6 presents the results of Eq. (5) for low-skilled rural migrant workers. The coefficients of individual-level control variables are significant and in line with the expectations of economic theories.

<< Insert Table 6 here >>

In the central region, the two coefficients of the interaction terms are both significant and positive which, according to Proposition 1, indicates that the increase of minimum wages has a significant and positive impact on employment of low-skilled rural migrant workers. As it is a nonlinear model, we further calculate the treatment effect of minimum wages following the method proposed by Puhani (2012).[11] According to Proposition 5, the estimated ATE on employment is 0.100 ($se = 0.016$,

[11] Our results are robust to the estimation of a linear probability model.

significant at 1% level). In the western region, the two DID estimators are both positive, but only one of them is significant. According to Proposition 3 and Proposition 5, it can be deduced that the employment effect is positive and the ATE is around 0.032 ($se = 0.016$, significant at 5% level).

Because Proposition 1 relies on less restrictive assumptions than Proposition 3, a more robust economic interpretation of the results above is that minimum wages have a significant and positive impact on employment of low-skilled rural migrant workers in the central region and may possibly have a positive impact on employment of low-skilled migrant workers in the western region. More importantly, there is no evidence that minimum wages have a negative impact on employment of rural migrant workers. These results stand in contrast to some findings based on data before 2010 (Ding, 2010; Fang and Lin, 2015), but are consistent with some recent evidence (Mayneris et al., 2018; Wang and Gunderson, 2018).

6.3. *Robustness check*

In this section, we assess the robustness of our results in several aspects. First, it is important to ensure the before-after comparability of repeated cross-sectional data, particularly given the high mobility of rural migrant workers. In our baseline results, we include individual characteristics as controls. To better control for group differences, we exploit the propensity score kernel matching DID estimator. Variables in X_{idt} are used to calculate the propensity score and matching is performed on three control groups (treated and nontreated in the first period and nontreated in the second period) to ensure both the before-after comparability and treat-control comparability (see Blundell and Dias (2009) for details). The results, which are similar to our previous findings, are presented in Table 7.

<< ***Insert Table 7 here*** >>

Another concern is that many rural migrant workers are self-employed who are not subject to minimum wages. Most minimum wages studies in developing countries find that minimum wages have no significant impact on self-employed groups (Lemos, 2009). Our previous analysis therefore follows the convention excluding self-employed workers. However, some recent studies claim that minimum wages may have an impact on self-employment, although without consensus (Yang and Zhang, 2020). According to the CMDS data, about 33.75% of rural migrant workers are self-employed. If the increase of minimum wages does have an impact on the self-employed group, our previous results will be biased.

Therefore, we include self-employed workers in the analysis of the effect of minimum wages on employment of low-skilled rural migrant workers.[12] The results are similar to our previous findings (see Table 8).

Finally, we replicate our research with the 2017 CMDS data. We match provinces groups following the same criterion used in the previous analysis. As most provinces did not adjust minimum wages in 2016, we obtain only two matched-provinces groups to apply the crossover design: the Hunan-Hubei group (2014-2016) and the Anhui-Jiangxi group (2012-2015). However, the Hunan-Hubei group is further excluded from the analysis because the results of the crossover design fail to detect significant wages effects.[13] Interestingly, Jiangxi raised the minimum wage between 2014 and 2015 while Anhui did not. Combing the previous analysis, we get a four periods crossover design during 2012-2015.

<< ***Insert Table 8 here*** >>

As shown in Table 9, the wages effects of minimum wages among low-skilled rural migrant workers are all positive and significant whereas the employment effect is positive but insignificant during 2014-2015. According to Proposition 5, the estimated ATE on employment during 2013-2015 is 0.063 ($se = 0.017$, significant at 1% level). We still find no evidence that minimum wages negatively affect employment.

<< ***Insert Table 9 here*** >>

6.4. *Discussion*

Our results show that raising minimum wages increases wages of low-skilled rural migrant workers in the central and western regions but does not adversely affect their employment. According to the modern monopsony model (Boal and Ransom, 1997; Ashenfelter, 2010), the results are valid under the

[12] A conventional assumption about self-employed workers is that they are displaced from the formal to informal sector. However, recent research shows that many self-employed migrant workers in China are self-selected rather than forced (Cui et al, 2015). In our sample, even for low-skilled workers, the average income of self-employed (3642.81 yuan) is significantly higher than that of employees (2799.30 yuan), contradicting the view of self-employment as an inferior choice. Therefore, there is no clear theory on how minimum wages impact self-employed in China. We analyzed the effect of minimum wages on self-employment of low-skilled rural migrants and the results are upon request.

[13] In addition, the wages effects in the high-skilled group ($\delta_A = 0.042, se = 0.027; \delta_B = 0.069, se = 0.031$) are higher than those in the low-skilled group ($\delta_A = -0.002, se = -0.024; \delta_B = 0.002, se = 0.030$), which is inconsistent with the expectation.

following two conditions: (i) in the low-skilled labor market, employers have stronger bargaining power than rural migrant workers; (ii) the minimum wage is higher than the market equilibrium wage but lower than the social optimal minimum wage (i.e., the "fair" wage in the competitive labor market).

Labor laws in China require employers to give workers labor contracts and pay social insurances premiums for full-time workers. According to the CMDS data we use, 55.71% of employed rural migrant workers have labor contracts, whereas the proportion with medical insurance coverage is only 28.47%. Using data from the Chinese National Bureau of Statistics, Liu and Kuruvilla (2017) also find that labor contracts and various social insurances are enforced poorly and unevenly among rural migrant workers. The weak and uneven enforcement of labor laws suggests that the bargaining power of rural migrant workers is weaker than that of employers. Therefore, the labor market of low-skilled rural migrant workers meets the first condition.

It is generally agreed that the minimum wage should be set at the level of 40% - 60% of the average social wage. Nevertheless, there is no official social average wage in China. Using the average wage of urban employees reported in the Chinese Statistical Yearbook as the measure of the social average wage would underestimate the relative level of the minimum wage (the ratio is 0.23 in 2013) because rural migrant workers and informal workers are not covered in the statistics. With employees in private firms and self-employed included, Ye et al. (2016) report that the ratio of the minimum wage to the social average wage is 0.30-0.33 during 2004-2013. If the overtime wage regulations were strictly implemented, the ratio of the minimum wage to the social average wage would increase to 0.43 in 2013 (Ye et al., 2016), which is comparable with the international norm. However, enforcement of the overtime regulations is very weak, especially for rural migrant workers (Ye et al., 2015). Therefore, from the perspective of international comparison, China's minimum wages are still low. In addition, Lu and Zheng (2019) find that the lower bargaining power of rural migrant workers has resulted in wages that are 24.96% lower than "fair" wages in the labor market. All of these findings suggest that the second condition also holds.

7. Conclusions

As a common method in the minimum wage literature, the DID design relies on the parallel trends assumption. However, it may not be feasible to check this assumption in many DID applications, leading to important divergence in empirical conclusions. In this study we propose a crossover design where one of the two study groups crosses over from the treatment group in the first period to the control group in the second period. We show that applying the crossover design researchers can obtain reliable inference on treatment effects without resting on the parallel trends assumption. The crossover design can also help identify spurious results.

Using individual data from the 2011-2014 CMDS, we employ the crossover design to estimate the wages and employment effects of minimum wages in the central, western, and northeast regions of China, respectively. We find robust evidence that in the central and eastern regions the increase of minimum wages has a significant and positive impact on wages of low-skilled rural migrant workers but an insignificant impact on wages of high-skilled workers. Our analysis does not find any evidence that minimum wages may adversely affect employment. Instead, the increase of minimum wages may have a positive effect on employment of low-skilled rural migrant workers. In the northeastern region, however, the parallel trends assumption of wages is violated and the wages and employment effect of minimum wages cannot be reliably identified. All in all, our results suggest that there may still be room to raise minimum wages in China without deceasing employment.

References

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Ashenfelter, Orley C., Henry Farber, and Michael R. Ransom. 2010. Labor market monopsony. *Journal of Labor Economics 28*(2): 203-210.

Athey, Susan, and Guido W. Imbens. 2022. Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics 226*(1): 62-79.

Bilinski, Alyssa, and Laura A. Hatfield. 2019. Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions. Working Paper, arXiv:1805.03273 [stat.ME],

Blundell, Richard and Monica Costa Dias. 2009. Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources 44*(3): 565-640.

Boal, William M., and Michael R. Ransom. 1997. Monopsony in the labor market. *Journal of economic literature 35*(1): 86-112.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2): 200-230.

Card, David, and Alan B. Krueger. 1994. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review 84*(4): 772-93.

Card, David, and Alan B. Krueger. 2000. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: reply. *American Economic Review 90*(5): 1397-1420.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134(*3): 1405-1454.

Cui, Yuling, Nahm Daehoon, and Tani Massimiliano. 2015. Self-employment in China: Are rural migrant workers and urban residents alike?. *Australian Economic Review 48(4)*: 382-399.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2018. Fuzzy differences-in-differences. *The Review of Economic Studies 85*(2): 999-1028.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9): 2964-96.

Ding, Shouhai. 2010. Employment effects of minimum wage regulation and cross effect of the employment contracts law. *Social Sciences in China 31*(3): 146-167.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. Minimum wage effects across state borders: Estimates using contiguous counties. *The Review of Economics and Statistics 92*(4): 945-964.

Dube, Arindrajit. 2019. Minimum wages and the distribution of family incomes. *American Economic Journal: Applied Economics 11*(4): 268-304.

Economides, George, and Thomas Moutos. 2016. Can minimum wages raise workers' incomes in the long run? *Journal of Public Economic Theory 18*(6): 961-978.

Fang, Tony, and Carl Lin. 2015. Minimum wages and employment in China. *IZA Journal of Labor Policy 4*(1): 1-30.

Fang, Tony, Morley Gunderson, and Carl Lin. 2021. The impact of minimum wages on wages, wage spillovers, and employment in China: Evidence from longitudinal individual-level data. *Review of Development Economics 25*(2): 854-877.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro. 2019. Pre-event trends in the panel event-study design. *American Economic Review 109*(9): 3307-38.

Goodman-Bacon, Andrew. 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2): 254-277.

Hoffman, Saul D. 2016. Are the effects of minimum wage increases always small? A reanalysis of Sabia, Burkhauser, and Hansen. *ILR Review 69*(2): 295-311.

Jia, Peng. 2014. Employment and working hour effects of minimum wage increase: Evidence from China. *China & World Economy 22*(2): 61-80.

Lemos, Sara. 2009. Minimum wage effects in a developing country. *Labour Economics 16*(2): 224-237.

Liu, Mingwei, and Sarosh Kuruvilla. 2017. The State, the Unions, and Collective Bargaining in China: The Good, the Bad, and the Ugly. *Comparative Labor Law and Policy Journal 38*(2): 187-210.

Liu, Mingwei. 2009. Chinese Employment Relations and Trade Unions in Transition. PhD Diss., School of Industrial and Labor Relations, Cornell University.

Lu, Haiyang, and Xuyuan Zheng. 2019. Difference in endowment, bargaining power and the wages of rural-urban migrants. *Journal of Agrotechnical Economics (6)*: 97-106 (in Chinese).

Maclure, Malcolm, and MA Mittleman. 2000. Should we use a case-crossover design? *Annual Review of Public Health 21*(1): 193-221.

Manning, Alan. 2013. *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Manning, Alan. 2021. The elusive employment effect of the minimum wage. *Journal of Economic Perspectives 35*(1): 3-26.

Manski, Charles F., and John V. Pepper. 2018. How do right-to-carry laws affect crime rates? Coping with ambiguity using bounded-variation assumptions. *Review of Economics and Statistics 100*(2): 232-244.

Mayneris, Florian, Sandra Poncet, and Tao Zhang. 2018. Improving or disappearing: Firm-level adjustments to minimum wages in China. *Journal of Development Economics*, *135*: 20-42.

Meer, Jonathan, and Jeremy West. 2016. Effects of the minimum wage on employment dynamics. *Journal of Human Resources 51*(2): 500-522.

Meyer, Breed D. 1995. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics 13*(2): 151-161.

Mora, Ricardo, and Iliana Reggio. 2019. Alternative diff-in-diffs estimators with several pretreatment periods. *Econometric Reviews 38*(5): 465-486.

Mora, Ricardo, and Iliana Reggio. 2019. Alternative diff-in-diffs estimators with several pretreatment periods. *Econometric Reviews* 38(5): 465-486.

Neumark, David, and Peter Shirley. 2021. Myth or measurement: What does the new minimum wage research say about minimum wages and job loss in the United States? NBER Working Paper 28388, National Bureau of Economic Research, Cambridge, MA.

Neumark, David, JM Ian Salas, and William Wascher. 2014. Revisiting the minimum wage—Employment debate: Throwing out the baby with the bathwater? *ILR Review 67*(3_suppl): 608-648.

Neumark, David, William L. Wascher, and William L. Wascher. 2008. *Minimum wages*. MIT press.

Ni, Jinlan, Guangxin Wang, and Xianguo Yao. 2011. Impact of minimum wages on employment: evidence from China. *Chinese Economy 44*(1): 18-38.

Puhani, Patrick A. 2012. The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters 115*(1): 85-87.

Raghavarao, Damaraju, and Lakshmi V. Padgett. 2014. *Repeated measurements and cross-over Designs*, John Wiley & Sons.

Rambachan, Ashesh, and Jonathan Roth. 2022. An honest approach to parallel trends. Unpublished manuscript, Department of Economics, Harvard University.

Sabia, Joseph J., Richard V. Burkhauser, and Benjamin Hansen. 2012. Are the effects of minimum wage increases always small? New evidence from a case study of New York state. *ILR Review 65*(2): 350-376.

Sabia, Joseph J., Richard V. Burkhauser, and Benjamin Hansen. 2016. When good measurement goes wrong: New evidence that New York state's minimum wage reduced employment. *ILR Review 69*(2): 312-319.

Sorkin, Isaac. 2015. Are there long-run effects of the minimum wage? *Review of Economic Dynamics 18*(2): 306-333.

Sun, Liyang, and Sarah Abraham. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2): 175-199.

Wang, Jing, and Morley Gunderson. 2012. Minimum Wage effects on employment and wages: diff-in-diff estimates from eastern China. *International Journal of Manpower 33*(8): 860-876.

Wang, Jing, and Morley Gunderson. 2018. Minimum wages effects on low-skilled workers in less developed regions of China. *International Journal of Manpower 39*(3): 455-467.

Yang, Juan, and Shi Li. 2016. Will the increase of minimum wages improve rural migrants' wages? *China Economic Quarterly 15*(4): 1563-1580 (in Chinese).

Yang, Zhengxiong, and Shiwei Zhang. 2020. The effects of minimum wages on informal employment choice and wages of rural migrant workers. *Issues in Agricultural Economy (9)*: 40-54 (in Chinese).

Ye, Linxiang, Shi Li, and Chan Yang. 2016. Estimating the relative level of minimum wage in China. *Chinese Journal of Population Science (01):*99-107+128 (in Chinese).

Ye, Linxiang, Thomas H. Gindling, and Shi Li. 2015. Compliance with legal minimum wages and overtime pay regulations in China. *IZA Journal of Labor & Development 4*(1): 1-35.

Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66 (5): 688.

Robins, James. 1986. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling* 7(9-12): 1393-1512.

Kahn-Lang, A., and Lang, K. 2020. The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics* 38(3): 613-620.

Roth, Jonathan. 2022. Pre-test with Caution: Event-study estimates after testing for parallel trends, forthcoming, *American Economic Review: Insights*

Roth, J., and Sant'Anna, P. H. 2022. Efficient estimation for staggered rollout designs. Working paper.

Table 1

Status of minimum wages adjustments across provinces, 2011-2014

	2012 (July 2011- May 2012)	2013 (May 2012-May 2013)	2014 (May 2013-May 2014)
Eastern region:			
Beijing, Tianjin, Shanghai, Shandong	1	1	1
Hebei	1	1	0
Jiangsu, Fujian, Hainan	0	1	1
Zhejiang	0	1	0
Guangdong	0	0	1
Central region:			
Shanxi, Hunan	1	1	1
Jiangxi, Henan	1	**1**	**0**
Anhui, Hubei	1	**0**	**1**
Western region:			
Shaanxi, Yunnan, Mongolia, Gansu	1	1	1
Guangxi, Guizhou, Qinghai	1	1	0
Sichuan, Ningxia	**1**	**0**	1
Chongqing, Xinjiang	**0**	**1**	1
Tibet	0	1	0
Northeastern region:			
Liaoning	**1**	**0**	1
Jilin	**0**	**1**	1
Heilongjiang	**0**	**1**	0

Notes: Shenzhen is not included in Guangdong Province. 1 indicates that a province adjusted minimum wages between the previous and current CMDS surveys, and 0 otherwise. The provinces and periods included in our crossover design are highlighted in black font.

Table 2

GDP per capita and minimum wages standards of the matched provinces（RMB）

	GDP Per Capita					Minimum Wage			
	2011	2012	2013	2014	AAGR	2011	2012	2013	2014
Central region:									
Jiangxi	26150	28800	31930	34661	0.098	720	870	1230	1230
Anhui	25659	28792	32001	34427	0.103	720	1010	1010	1260
Western region:									
Chengdu	48755	56836	64248	70338	0.130	850	1050	1050	1200
Chongqing	34500	38914	43223	47859	0.115	870	870	1050	1250
Northeastern region:									
Liaoning	50760	56649	61996	65201	0.087	900	1100	1100	1300
Jilin	38460	43415	47428	50162	0.093	1000	1000	1150	1320
Heilongjiang	32819	35711	37697	39226	0.061	880	880	1160	1160

Notes: AAGR denotes average annual growth rate of GDP per capita. First-tier minimum wages standards on the survey date are presented.

Table 3

Individual characteristics of rural migrant workers (by provinces and skill groups)

	Monthly wages		Employment	Education		Age		Male	Married	*N*
	Mean	SD	Mean	Mean	SD	Mean	SD	Mean	Mean	
High-skilled Group:										
Central region:										
Jiangxi	2438.431	1062.536	0.897	12.724	1.374	27.132	7.274	0.541	0.396	2610
Anhui	2979.943	1420.656	0.750	12.85	1.434	28.487	6.576	0.557	0.762	1991
p-value	[0.000]		[0.000]	[0.003]		[0.000]		[0.283]	[0.000]	
Western region:										
Chengdu	2522.729	1109.708	0.940	12.705	1.34	27.381	7.304	0.524	0.429	1411
Chongqing	2738.410	1176.747	0.909	12.937	1.514	28.176	7.857	0.559	0.473	3156
p-value	[0.000]		[0.001]	[0.000]		[0.001]		[0.029]	[0.006]	
Northeastern region:										
Liaoning	2731.814	1143.918	0.902	12.769	1.438	29.369	7.891	0.537	0.425	2439
J&H	2615.152	1162.127	0.790	12.776	1.428	30.341	8.917	0.529	0.54	2814
p-value	[0.001]		[0.000]	[0.854]		[0.000]		[0.564]	[0.000]	
Low-skilled Group:										
Central region:										
Jiangxi	2359.436	1017.443	0.802	8.380	1.588	33.255	9.582	0.449	0.736	4704
Anhui	2584.032	1316.577	0.639	8.242	1.798	34.267	7.726	0.395	0.947	5705
p-value	[0.000]		[0.000]	[0.000]		[0.000]		[0.000]	[0.000]	
Western region:										
Chengdu	2300.096	1007.763	0.921	8.316	1.471	35.391	9.941	0.51	0.779	2706
Chongqing	2458.719	1150.342	0.851	8.205	1.579	37.203	10.383	0.529	0.816	5907
p-value	[0.000]		[0.000]	[0.002]		[0.000]		[0.088]	[0.000]	
Northeastern region:										
Liaoning	2594.724	1144.645	0.865	8.521	1.285	33.539	10.245	0.542	0.637	7422
J&H	2437.232	1154.940	0.676	8.282	1.583	36.836	10.332	0.504	0.826	13576
p-value	[0.000]		[0.000]	[0.000]		[0.000]		[0.000]	[0.000]	

Sources: CMDS (2011-2014).

Notes: p-values for differences in mean values of the individual characteristics are reported in square brackets.

Table 4
Wage equations of high-skilled rural migrant workers

	Jiangxi-Anhui		Chengdu-Chongqing		Liaoning- J&H	
	Year 12-13	Year 13-14	Year 11-12	Year 12-13	Year 11-12	Year 12-13
Edu	0.046	0.037	0.050	0.056	0.048	0.038
	(0.007)	(0.005)	(0.007)	(0.006)	(0.007)	(0.006)
Exp	0.028	0.034	0.024	0.027	0.025	0.028
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Exp^2	-0.082	-0.100	-0.096	-0.090	-0.074	-0.080
	(0.012)	(0.011)	(0.014)	(0.011)	(0.011)	(0.009)
Male	0.182	0.204	0.157	0.156	0.162	0.104
	(0.021)	(0.019)	(0.023)	(0.019)	(0.024)	(0.018)
Married	-0.063	-0.050	0.037	-0.010	-0.120	-0.131
	(0.031)	(0.026)	(0.033)	(0.026)	(0.030)	(0.024)
Male * Married	0.252	0.226	0.154	0.160	0.198	0.252
	(0.033)	(0.028)	(0.037)	(0.031)	(0.037)	(0.029)
Post	0.121	0.120	0.096	0.172	0.185	0.052
	(0.029)	(0.017)	(0.023)	(0.028)	(0.032)	(0.020)
Treat * Post (Treatment effect)	0.021	-0.021	0.032	-0.027	-0.073	0.049
	(0.035)	(0.030)	(0.036)	(0.034)	(0.039)	(0.030)
Constant	6.621	6.928	6.747	6.732	6.659	7.089
	(0.263)	(0.091)	(0.102)	(0.081)	(0.103)	(0.080)
City fixed effect	Yes	Yes	—	—	Yes	Yes
N	1935	2324	1496	2076	1567	2206
R^2	0.320	0.332	0.214	0.229	0.204	0.176

Notes: Results are from the estimation of Eq. (4). Robust standard errors are reported in parentheses.

Table 5
Wage equations of low-skilled rural migrant workers

	Jiangxi-Anhui		Chengdu-Chongqing		Liaoning-J&H	
	Year 12-13	Year 13-14	Year 11-12	Year 12-13	Year 11-12	Year 12-13
Edu	0.007	0.004	0.004	0.006	-0.004	0.016
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Exp	0.014	0.013	0.018	0.018	0.022	0.019
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Exp^2	-0.044	-0.040	-0.049	-0.046	-0.057	-0.051
	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)
Male	0.117	0.088	0.135	0.151	0.184	0.136
	(0.024)	(0.023)	(0.028)	(0.024)	(0.015)	(0.013)
Married	-0.068	-0.082	-0.034	-0.056	-0.121	-0.090
	(0.026)	(0.026)	(0.032)	(0.027)	(0.020)	(0.016)
Male * Married	0.404	0.409	0.254	0.237	0.278	0.305
	(0.028)	(0.026)	(0.032)	(0.027)	(0.020)	(0.017)
Post	0.111	0.093	0.069	0.129	0.052	0.121
	(0.018)	(0.015)	(0.016)	(0.019)	(0.014)	(0.012)
Treat * Post (Treatment effect)	0.048	0.043	0.044	0.085	0.049	0.000
	(0.025)	(0.023)	(0.025)	(0.024)	(0.019)	(0.017)
Constant	7.418	7.414	7.208	7.287	7.267	7.326
	(0.126)	(0.059)	(0.049)	(0.049)	(0.043)	(0.043)
City fixed effect	Yes	Yes	—	—	Yes	Yes
N	3782	4147	3284	3822	6101	7402
R^2	0.344	0.339	0.228	0.264	0.242	0.248

Notes: Results are from the estimation of Eq. (4). Robust standard errors are reported in parentheses.

Table 6

Employment equations of low-skilled rural migrant workers

	Jiangxi-Anhui		Chengdu-Chongqing	
	Year 12-13	Year 13-14	Year 11-12	Year 12-13
Edu	0.035	0.011	0.035	0.050
	(0.014)	(0.013)	(0.016)	(0.017)
Exp	0.152	0.139	0.202	0.193
	(0.025)	(0.022)	(0.023)	(0.024)
Exp^2	-0.192	-0.173	-0.262	-0.245
	(0.036)	(0.030)	(0.030)	(0.031)
Male	-0.335	-0.252	0.428	0.034
	(0.158)	(0.125)	(0.176)	(0.155)
Married	-1.974	-1.572	-1.451	-1.467
	(0.155)	(0.131)	(0.167)	(0.163)
Male * Married	2.406	1.881	0.948	1.251
	(0.177)	(0.138)	(0.192)	(0.173)
Post	-0.105	-0.322	0.165	-0.026
	(0.057)	(0.071)	(0.065)	(0.112)
Treat * Post	0.475	0.398	0.291	0.129
	(0.095)	(0.088)	(0.125)	(0.129)
Constant	-1.158	-0.721	-2.179	-1.457
	(0.447)	(0.390)	(0.409)	(0.416)
City fixed effect	Yes	Yes	—	—
N	5247	5534	3889	4332
Pseudo R^2	0.325	0.258	0.207	0.177
Treatment effect	0.101	0.099	0.040	0.024
	(0.021)	(0.024)	(0.018)	(0.025)

Notes: Results are from the estimation of Eq. (5). Treatment effects are calculated by Eq. (6). Robust standard errors are reported in parentheses.

Table 7

Minimum wages effects on low-skilled rural migrant workers: PSM+DID

	Jiangxi-Anhui		Chengdu-Chongqing	
	Year 12-13	Year 13-14	Year 11-12	Year 12-13
Wages effect	0.070	0.052	0.058	0.090
	(0.035)	(0.030)	(0.028)	(0.028)
Employment effect	0.087	0.057	0.039	0.029
	(0.028)	(0.025)	(0.022)	(0.019)

Notes: Results are from the estimation of the PSM+DID equation. The estimated treatment effects of minimum wages are reported. Robust standard errors are reported in parentheses.

Table 8

Minimum wages effect on employment of low-skilled rural migrant workers: self-employed included

	Jiangxi-Anhui		Chengdu-Chongqing	
	Year 12-13	Year 13-14	Year 11-12	Year 12-13
Treatment effect	0.088	0.065	0.027	0.025
	(0.014)	(0.017)	(0.011)	(0.020)
Controls	Yes	Yes	Yes	Yes
N	9399	10381	6046	6416
Pseudo R^2	0.280	0.227	0.183	0.168

Notes: Results are from the estimation of Eq. (4) and Eq. (5). Treatment effect on employment is calculated by Eq. (6). Robust standard errors are reported in parentheses.

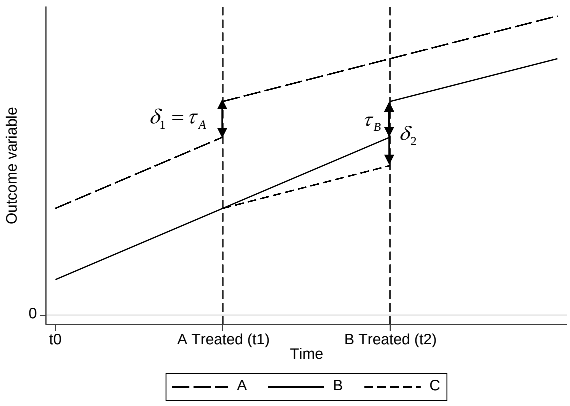
Table 9

Minimum wages effects on low-skilled rural migrant workers: Jiangxi-Anhui group, 2012-2015

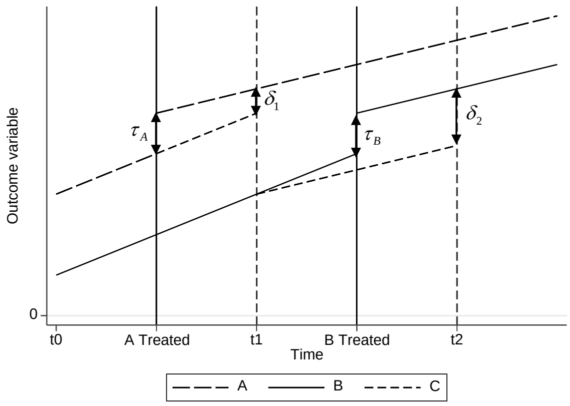
	Year 12-13	Year 13-14	Year 14-15
Wage effect	0.048	0.043	0.039
	(0.025)	(0.023)	(0.023)
Employment effect	0.101	0.099	0.026
	(0.021)	(0.024)	(0.023)

Notes: Results are from the estimation of Eq. (4) and Eq. (5). Treatment effect on employment is calculated by Eq. (6). Robust standard errors are reported in parentheses.

A. The treatment has both an instant and a dynamic impact on the outcome variable

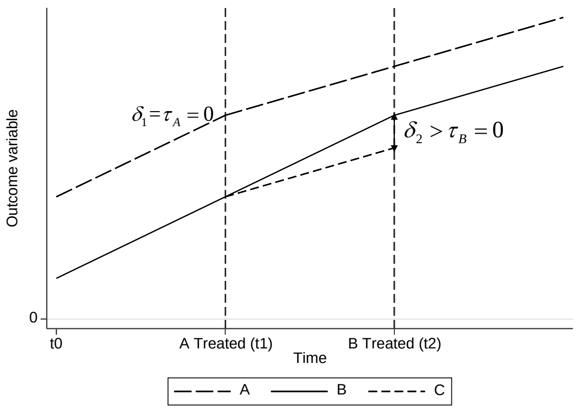


(a)

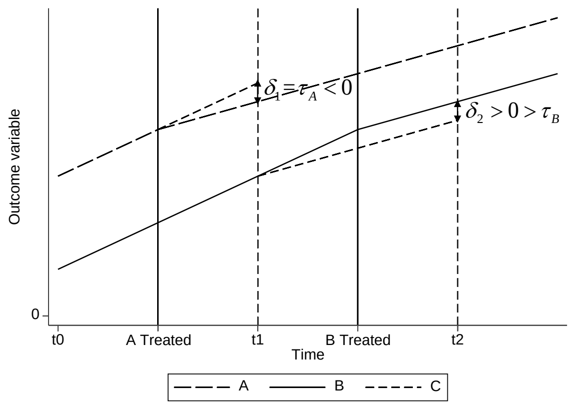


(b)

B. The treatment only impacts the growth rate of the outcome variable



(c)



(d)

Figure 1

Two cases of violations of Assumption 4a

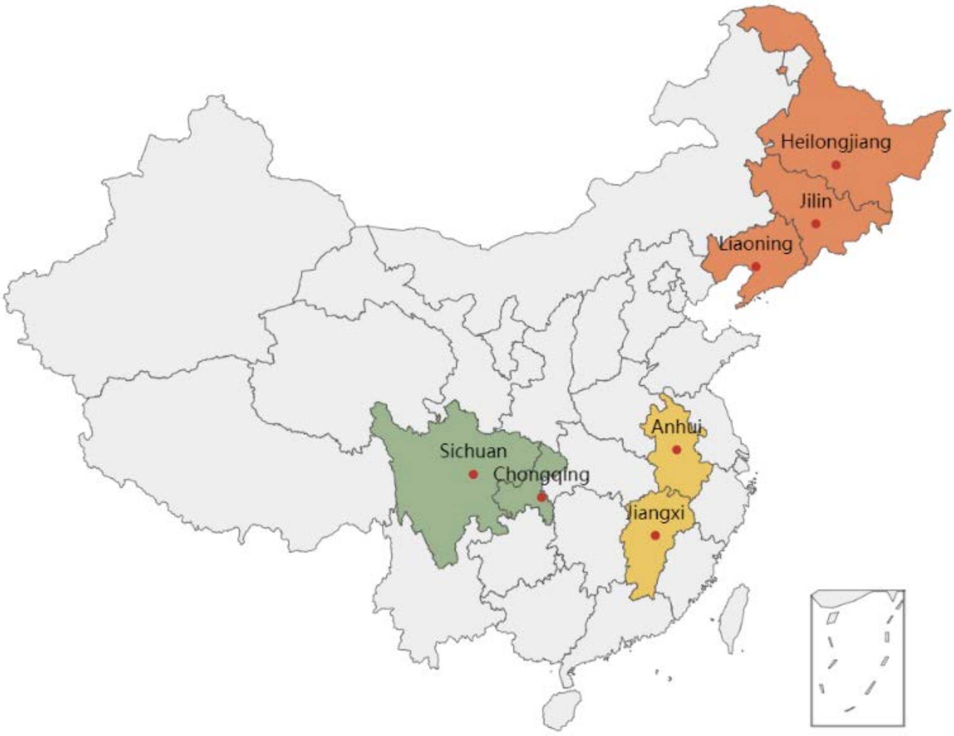


Figure 2
Matched provinces in the crossover design