BIOMETRICS 64, 1–?? DOI: 000
December 2008

Discussion on "Instrumented difference-in-differences" by Ting Ye, Ashkan Ertefaie, James Flory, Sean Hennessy, and Dylan S. Small

Hyunseung Kang*

Department of Statistics, University of Wisconsin-Madison, 1300 University Ave., Madison 53706, U.S.A.

**email:* hyunseung@stat.wisc.edu

SUMMARY: We reinterpret the instrumented difference-in-differences (iDID) under a linear instrumental variables (IV) model. Under the linear IV model, we show why iDID is a clear improvement over two existing methods, difference-in-differences (DID) and a cross-sectional, IV analysis. We also re-express some of the assumptions of iDID using familiar, regression-based identification assumptions. We conclude with a method inspired by the linear IV model that can potentially remedy the weak identification problem in iDID.

KEY WORDS: Anderson-Rubin test; Difference-in-differences; Instrumental variables; Linear models; Weak identification

1. Introduction

I want to congratulate the authors for their fantastic work on combining two well-known methods in causal inference, difference-in-differences (DID) and instrumental variables (IV), in order to study causal exposure effects in repeated, cross-sectional observational studies. By combining the strengths from each method, the proposed instrumented difference-in-differences (iDID) is more robust to violations of the DID's parallel trend assumption due to an unmeasured confounder and the IV's exclusion restriction where iDID can use an instrument that has a direct effect on the outcome.

The goal of the paper is to reinterpret this promising method under a simple, but popular model in econometrics and statistics, a linear IV model. Linear IV models (or linear models, in general) have been the workhorse in applied statistics and economics to conduct data analysis. Also, in non-applied works, linear IV models have been used to build theoretical insights and create more robust, less parametric methods. In fact, most textbook introductions of IV or DID in econometrics use linear models as a "reference" model to ground key ideas and discuss more complex topics (e.g. Chapter 5 of Angrist and Pischke (2008) or Chapter 5 of Wooldridge (2010)). By using a linear IV model, I wish to provide an alternative explanation of the authors' fantastic method that is (hopefully) more familiar, simple, and accessible.

The paper will primarily focus on three aspects of iDID under the linear IV model:

(a) why iDID is a clear improvement over DID or a cross-sectional, IV analysis;

(b) how some of the identifying assumptions of iDID can be re-expressed using traditional, regression-based assumptions;

(c) how to use insights from linear IV models to potentially mitigate the weak identification problem discussed in Section 5 of the authors' work.

Of course, according to George Box's famous aphorism, all models are wrong and the linear IV model used in the paper, while based on the authors' results (see Section 2.1), is no exception.

However, I hope the model is still useful, especially for investigators contemplating to use iDID in their observational studies.

2. Reframing the Problem With a Linear IV Model

2.1 *Review of Section 5 and Setup*

We first review Section 5 of the authors' work where the connection between iDID and a linear IV model is hinted from a result concerning the properties of one of the authors' proposed estimator, $\widehat{\beta}_{\text{wald}}$. Formally, in the absence of covariates, consider the following model for individual i's observed data $\mathbf{O}_i = (Y_i, D_i, Z_i, T_i)$ where, identical to the authors' notation, Y_i is a real-valued outcome, D_i is a binary exposure, Z_i is a binary instrument, and T_i is a binary time indicator:

$$Y_i = \beta_{\text{int}} + \beta_{\text{D}} D_i + \beta_{\text{Z}} Z_i + \beta_{\text{T}} T_i + \epsilon_i, \quad E[\epsilon_i \mid Z_i, T_i] = 0. \tag{1}$$

The terms $\beta_{\text{int}}, \beta_{\text{D}}, \beta_{\text{Z}}, \beta_{\text{T}}$ are unknown parameters of the model and the term ϵ_i is a random error term that has mean zero given the regressors Z_i and T_i, but not the regressor D_i. Using econometrics terminology, Z_i and T_i are exogenous regressors (i.e. independent from the error term) and D_i is an endogenous regressor (i.e. dependent on the error term). The authors showed that the well-known, two-stage least squares (2SLS) estimator of β_{D} with an "interacted" instrument $Z_i T_i$ is numerically equivalent to one of their proposed estimators, $\widehat{\beta}_{\text{wald}}$. That is, in the first step, we regress D_i on the intercept, Z_i, T_i, and $Z_i T_i$ and obtain the predicted value of D_i, denoted as $\widehat{D}_i$; note that the regression in the first step must be linear. In the second step, we regress Y_i on the intercept, Z_i, T_i, and $\widehat{D}_i$ and the authors showed that the estimated coefficient for the regressor $\widehat{D}_i$ (i.e. the 2SLS estimator of β_{D}) is numerically equivalent to $\widehat{\beta}_{\text{wald}}$.

We make a few remarks about model (1) that may be useful for extending iDID to other data types. First, if covariates $\mathbf{X}_i$ are present, we can incorporate them in model (1), say by adding a linear term $\boldsymbol{\beta}_{\mathrm{X}}^{\intercal}\mathbf{X}_i$ where $\boldsymbol{\beta}_{\mathrm{X}}$ is another unknown parameter whose dimension is equal to the dimension of $\mathbf{X}_i$. But, this introduces additional modeling assumptions about the relationship between $\mathbf{X}_i$ and Y_i. Second, if there are multiple time points T_i or if the instrument Z_i is non-binary, model (1) provides one simple, starting point to extend iDID, especially its estimation framework. For example, for multiple time points, investigators can represent time as fixed effects in model (1) and use one (or several) interacted instruments between Z_i and each level of T_i. For a non-binary instrument, model (1) can be used as-is or can be modified to reflect the instrument's potentially non-linear effect on the outcome.

Taking inspiration from the authors' numerical equivalence result, the rest of the paper will assume that model (1) is the true model for the observed data. But, as forewarned in Section 1, if the model is misspecified, the discussion below may be dangerously misleading and readers should consult the authors' work, which does not rely on a parametric model.

2.2 *Advantages of iDID versus DID or IV with model* (1)

Taking a step back from the authors' result on the equivalence between $\widehat{\beta}_{\mathrm{wald}}$ and the 2SLS estimator of β_{D}, the structure of the linear model (1) already reveals some of the advantages of iDID compared to DID or a cross-sectional, IV analysis. For example, under the usual DID setup without an instrument, the conditional mean of the error term ϵ_i given the exposure D_i and the time indicator T_i would be zero and consequently, the parallel trend assumption would hold; in other words, the usual DID setup assumes that the exposure D_i is exogenous. Instead, iDID allows the exposure to be endogenous and the parallel trend assumption may be violated due to an unmeasured confounder that affects the exposure and the outcome.

To better illustrate this point, consider a simple, hypothetical setup where we evaluate the usual DID estimator (denoted as $\widehat{\beta}_{\mathrm{DID}}$ and defined below) under model (1) where the exposure effect is zero (i.e. $\beta_{\mathrm{D}} = 0$) and the instrument satisfies the exclusion restriction (i.e. the instrument has no direct effect on the outcome so that $\beta_{\mathrm{Z}} = 0$). This exercise mimics an investigator who may initially run a DID analysis and assume that the exposure is exogenous, even though in reality, the exposure is endogenous due to unmeasured confounding. After some algebra, we get:

$$\begin{aligned}
\widehat{\beta}_{\mathrm{DID}} &\equiv \underbrace{\left(\frac{\sum_{i=1}^{n} Y_i D_i T_i}{\sum_{i=1}^{n} D_i T_i} - \frac{\sum_{i=1}^{n} Y_i (1-D_i) T_i}{\sum_{i=1}^{n} (1-D_i) T_i}\right)}_{\text{Average Exposure Effect at Time } T_i = 1} - \underbrace{\left(\frac{\sum_{i=1}^{n} Y_i D_i (1-T_i)}{\sum_{i=1}^{n} D_i (1-T_i)} - \frac{\sum_{i=1}^{n} Y_i (1-D_i)(1-T_i)}{\sum_{i=1}^{n} (1-D_i)(1-T_i)}\right)}_{\text{Average Exposure Effect at Time } T_i = 0} \\
&= \left(\frac{\sum_{i=1}^{n} (\beta_{\mathrm{T}} + \epsilon_i) D_i T_i}{\sum_{i=1}^{n} D_i T_i} - \frac{\sum_{i=1}^{n} (\beta_{\mathrm{T}} + \epsilon_i)(1-D_i) T_i}{\sum_{i=1}^{n} (1-D_i) T_i}\right) - \left(\frac{\sum_{i=1}^{n} \epsilon_i D_i (1-T_i)}{\sum_{i=1}^{n} D_i (1-T_i)} - \frac{\sum_{i=1}^{n} \epsilon_i (1-D_i)(1-T_i)}{\sum_{i=1}^{n} (1-D_i)(1-T_i)}\right) \\
&\to \underbrace{(E[\epsilon_i \mid D_i = 1, T_i = 1] - E[\epsilon_i \mid D_i = 0, T_i = 1])}_{\Delta_{T=1}} - \underbrace{(E[\epsilon_i \mid D_i = 1, T_i = 0] - E[\epsilon_i \mid D_i = 0, T_i = 0])}_{\Delta_{T=0}}.
\end{aligned}$$

The right arrow above represents the probability limit as the sample size goes to infinity and the limiting value is derived by using the law of large numbers. Roughly speaking, the term $\Delta_{T=1}$ represents the effect of unmeasured confounding at time $T = 1$ and $\Delta_{T=0}$ represents the effect of unmeasured confounding at time $T = 0$. If there is no unmeasured confounding at each time point and there are no covariates, the exposure is effectively randomly assigned to everyone at each time point, akin to running a randomized experiment at each time point, and the means of the error terms ϵ_i between the exposed (i.e. $D_i = 1$) and the unexposed (i.e. $D_i = 0$) groups would be the same, leading to $\Delta_{T=1} = 0$ and $\Delta_{T=0} = 0$. In other words, the usual DID estimator $\widehat{\beta}_{\mathrm{DID}}$ will converge to 0, as expected from this hypothetical setup. More generally, if the effect from unmeasured confounders are "identical" in magnitude at each time point where $\Delta_{T=1} = \Delta_{T=0}$, the DID estimator will still converge to 0; note that the parallel trend assumption implies $\Delta_{T=1} = \Delta_{T=0}$. However, if unmeasured confounders have different effects across time so that $\Delta_{T=1} \neq \Delta_{T=0}$, the parallel trend assumption is violated and $\widehat{\beta}_{\mathrm{DID}}$ no longer converges to 0.

Also, compared to a standard, cross-sectional IV analysis, iDID allows an instrument to violate the exclusion restriction. This can be clearly seen in model (1) where after fixing a particular time point $T_i = t$, the instrument Z_i can have a non-zero direct effect on the outcome Y_i through the term $\beta_{\mathrm{Z}} Z_i$. Also, if an investigator naively computes the usual Wald estimator in IV at $T_i = 0$ (denoted as $\widehat{\beta}_{\mathrm{wald},T=0}$ and defined below), the Wald estimator would evaluate to the following under model (1):

$$\begin{aligned}
\widehat{\beta}_{\text{wald},T=0} &\equiv \frac{\frac{\sum_{i=1}^{n} Y_i(1-T_i)Z_i}{\sum_{i=1}^{n}(1-T_i)Z_i} - \frac{\sum_{i=1}^{n} Y_i(1-T_i)(1-Z_i)}{\sum_{i=1}^{n}(1-T_i)(1-Z_i)}}{\frac{\sum_{i=1}^{n} D_i(1-T_i)Z_i}{\sum_{i=1}^{n}(1-T_i)Z_i} - \frac{\sum_{i=1}^{n} D_i(1-T_i)(1-Z_i)}{\sum_{i=1}^{n}(1-T_i)(1-Z_i)}} \\
&= \frac{\left(\beta_{\text{int}} + \beta_{\text{Z}} + \frac{\sum_{i=1}^{n}(\beta_{\text{D}} D_i+\epsilon_i)(1-T_i)Z_i}{\sum_{i=1}^{n}(1-T_i)Z_i}\right) - \left(\beta_{\text{int}} + \frac{\sum_{i=1}^{n}(\beta_{\text{D}} D_i+\epsilon_i)(1-T_i)(1-Z_i)}{\sum_{i=1}^{n}(1-T_i)(1-Z_i)}\right)}{\frac{\sum_{i=1}^{n} D_i(1-T_i)Z_i}{\sum_{i=1}^{n}(1-T_i)Z_i} - \frac{\sum_{i=1}^{n} D_i(1-T_i)(1-Z_i)}{\sum_{i=1}^{n}(1-T_i)(1-Z_i)}} \\
&\to \beta_{\text{D}} + \frac{\beta_{\text{Z}}}{E[D_i \mid T_i = 0, Z_i = 1] - E[D_i \mid T_i = 0, Z_i = 0]}.
\end{aligned}$$

The estimator $\widehat{\beta}_{\text{wald},T=0}$ is inconsistent for β_{D} unless the instrument satisfies the exclusion restriction by setting $\beta_{\text{Z}} = 0$. Alternatively, by having one additional sample at $T_i = 1$, a time-invariant instrument, and other assumptions stated in the authors' work, we can remove the bias arising from violating the exclusion restriction and consistently estimate β_{D}. Note that this is not the only way to consistently estimate β_{D} when the exclusion restriction is violated; see Kang et al. (2016), Guo et al. (2018), Tchetgen Tchetgen et al. (2021), and Ye et al. (2021) for some examples.

2.3 *Reinterpreting iDID assumptions with regression-based assumptions in linear IV models*

We can also use the well-established identifying conditions for model parameters in linear IV models, specifically a necessary condition known as the order condition (see Chapter 5.2.1 of Wooldridge (2010)), to reinterpret some of the identifying assumptions of iDID. To review, in linear IV models, the order condition roughly states that if the parameters in a linear IV model are identifiable, the number of instruments must be greater than or equal to the number of endogenous variables. In model (1), the order condition is satisfied because there is one instrument (i.e. $Z_i T_i$) and one endogenous variable (i.e. D_i).

Now suppose there is an interaction term between the exposure D_i and the time indicator T_i in model (1). If included, the interaction term would allow the effect of the exposure on the outcome (i.e. the exposure effect) to vary across time. But, including the interaction term would violate the order condition because there are more endogenous variables (i.e. D_i and

D_iT_i) than the number of instruments (i.e. Z_iT_i) and subsequently, the model parameters in model (1) are not identifiable. In the authors' work, Assumption (2d) is the "most relevant, nonparametric formulation" of this condition where the exposure effect is assumed to be homogeneous across time; here, we put the phrase "most relevant, nonparametric formulation" in quotes because formally tying model (1), the order condition, and the authors" nonparametric, identifying assumptions implicitly requires other assumptions in the authors' work, notably Assumption 1; see Section 4.4 of Holland (1988) for an example.

Similarly, suppose there is an interaction term between the exposure D_i and the instrument Z_i in model (1). If included, the interaction term would allow the exposure effect to vary between the encouraged (i.e. $Z_i = 1$) and the non-encouraged (i.e. $Z_i = 0$) groups. But, similar to the previous paragraph, including the interaction term would would violate the order condition. Also, in the authors' work, Assumption (2b) is the most relevant, nonparametric expression of this condition where the exposure effect is independent of the instrument Z_i. More generally, it's likely that most of the identifying assumptions of iDID are nonparametric extensions of the identification conditions for model parameters in a linear IV model.

2.4 *A Robust Confidence Interval Under Weak Identification: The Anderson-Rubin Interval*

Finally, we can use a simple, well-known method associated with linear IV models to potentially address the weak identification problem in iDID. To review, under Assumption (2a) in the authors' work, iDID requires an instrument that, on average, changes the trend in the exposure. But, when the instrument induces little to no change in the exposure's trend, the proposed point estimators may be biased and non-normal, a problem the authors refer to as the weak identification problem. The authors propose a diagnostic test to check for this problem by using a well-known F test for instrument strength in linear IV models where if the F test is sufficiently large, the proposed estimators may be less prone to bias.

In a similar vein, we can use a method inspired by linear IV models, specifically the work by Anderson and Rubin (1949), to come up with a valid $1-\alpha$ confidence interval that does not suffer from the weak identification problem. To motivate the confidence interval, suppose we want to test the null hypothesis that the coefficient β_{D} in model (1) is some hypothesized value $\beta_{D,0}$, i.e. $H_0 : \beta_{\mathrm{D}} = \beta_{D,0}$. After subtracting $D_i\beta_{D,0}$ from both sides of the equality in model (1) and taking expectations given Z_i and T_i, we arrive at

$$E[Y_i - D_i\beta_{D,0} \mid Z_i, T_i] = \beta_{\mathrm{int}} + (\beta_{\mathrm{D}} - \beta_{D,0})E[D_i \mid Z_i, T_i] + \beta_{\mathrm{Z}} Z_i + \beta_{\mathrm{T}} T_i. \tag{2}$$

Now, consider the following "model" for the conditional distribution of D_i given Z_i and T_i,

$$P(D_i = 1 \mid Z_i, T_i) = \gamma_{\mathrm{int}} + \gamma_{\mathrm{Z}} Z_i + \gamma_{\mathrm{T}} T_i + \gamma_{\mathrm{ZT}} Z_i T_i, \tag{3}$$

where the four terms $\gamma_{\mathrm{int}}, \gamma_{\mathrm{Z}}, \gamma_{\mathrm{T}}, \gamma_{\mathrm{ZT}}$ are unknown parameters. We put "model" in quotes because every conditional distribution of D_i given Z_i and T_i can be characterized by (3); in short, unlike (1), (3) is a saturated model of D_i given binary Z_i and T_i. Then, (2) becomes

$$\begin{aligned} E[Y_i - D_i\beta_{D,0} \mid Z_i, T_i] &= \{\beta_{\mathrm{int}} + (\beta_{\mathrm{D}} - \beta_{D,0})\gamma_{\mathrm{int}}\} \\ &\quad + \{\beta_{\mathrm{Z}} + (\beta_{\mathrm{D}} - \beta_{D,0})\gamma_{\mathrm{Z}}\} Z_i + \{\beta_{\mathrm{T}} + (\beta_{\mathrm{D}} - \beta_{D,0})\gamma_{\mathrm{T}}\} T_i + \underbrace{\{(\beta_{\mathrm{D}} - \beta_{D,0})\gamma_{\mathrm{ZT}}\}}_{\pi_{\mathrm{ZT}}} Z_i T_i \end{aligned} \tag{4}$$

Notice that equation (4) is a linear regression model with an "adjusted" outcome $Y_i - D_i\beta_{D,0}$ and regressors Z_i, T_i, and Z_iT_i. Thus, we can use ordinary least squares (OLS) to arrive at consistent estimators and/or tests of the parameters in the curly brackets above. Second, under the null $H_0 : \beta_{\mathrm{D}} = \beta_{D,0}$, the coefficient in front of the interaction term Z_iT_i in (4) (i.e. π_{ZT}) is zero, implying another null hypothesis $H_0 : \pi_{\mathrm{ZT}} = 0$. Critically, we can test the latter null hypothesis by using the usual (two-sided) t-test from the OLS estimate of π_{ZT} and its null distribution does not depend on how strong the instrument changes the trend in the exposure i.e. the term γ_{ZT} in (3).

The connection between testing $H_0 : \beta_{\mathrm{D}} = \beta_{D,0}$ and testing a regression coefficient is the basis for the Anderson-Rubin confidence interval for β_{D}. Formally, for a level $\alpha \in (0,1)$, we can test the regression coefficient $H_0 : \pi_{\mathrm{ZT}} = 0$ across different values of $\beta_{\mathrm{D},0}$ with the two-sided, t-test from OLS regression and by the duality between testing and confidence intervals, the accepted values of $\beta_{\mathrm{D},0}$ (i.e. the values of $\beta_{\mathrm{D},0}$ where $H_0 : \pi_{\mathrm{ZT}} = 0$ is accepted at level α) form a two-sided $1-\alpha$ confidence interval for β_{D}. These accepted values of $\beta_{\mathrm{D},0}$, denoted as $\mathcal{C}_{1-\alpha}^{\mathrm{AR}}$, can be compactly expressed as the following set

$$\mathcal{C}_{1-\alpha}^{\mathrm{AR}} = \left\{\beta_{\mathrm{D},0} \in \mathbb{R} \mid \frac{(\mathbf{Y}-\mathbf{D}\beta_{\mathrm{D},0})^{\intercal}\mathbf{H}_R(\mathbf{Y}-\mathbf{D}\beta_{\mathrm{D},0})}{(\mathbf{Y}-\mathbf{D}\beta_{\mathrm{D},0})^{\intercal}(\mathbf{I}-\mathbf{H}_R)(\mathbf{Y}-\mathbf{D}\beta_{\mathrm{D},0})/(n-4)} \leqslant \chi^2_{1-\alpha,1}\right\}, \qquad (5)$$

where, using the authors' notation, $\mathbf{Y}^{\intercal} = (Y_1,\ldots,Y_n)$, $\mathbf{D}^{\intercal} = (D_1,\ldots,D_n)$, $\mathbf{H}_R = \mathbf{R}(\mathbf{R}^{\intercal}\mathbf{R})^{-1}\mathbf{R}^{\intercal}$, and $\mathbf{R}$ is the n-dimensional vector of residuals from regressing Z_iT_i on the intercept, Z_i, and T_i. Also, $\chi^2_{1-\alpha,1}$ is the $1-\alpha$ quantile of the chi-square distribution with one degree of freedom. One of the most appealing properties of $C_{1-\alpha}^{\mathrm{AR}}$ is that compared to the Wald-based confidence interval in the authors' work, $\mathcal{C}_{1-\alpha}^{\mathrm{AR}}$ will always have at least $1-\alpha$ coverage irrespective of the instrument's association to the trend in the exposure. In the extreme case where Assumption (2a) is violated so that the target parameter is no longer point identified, $\mathcal{C}_{1-\alpha}^{\mathrm{AR}}$ will still have coverage by elongating itself to cover the entire real line, i.e. $\mathcal{C}_{1-\alpha}^{\mathrm{AR}} = (-\infty,\infty)$. While an infinite confidence interval may initially be unappealing, it alerts investigators about the lack of point identifiability from the observed data. Also, page 1377 of Dufour (1997) showed that a valid $1-\alpha$ confidence interval of β_{D} must be unbounded with non-zero probability and pages 133 and 134 of Moreira (2009) showed that under some assumptions, the test statistic underlying $\mathcal{C}_{1-\alpha}^{\mathrm{AR}}$ is the uniformly most powerful unbiased test for $H_0 : \beta_{\mathrm{D}} = \beta_{\mathrm{D},0}$.

Of course, there are no theoretical justifications for $C_{1-\alpha}^{\mathrm{AR}}$ outside of linear IV models. But, $C_{1-\alpha}^{\mathrm{AR}}$ could be a promising starting point to address the weak identification problem in iDID.

3. Final Thoughts

While the linear IV model (1) is undoubtedly too simple for real data and prone to misspecification, I hope the small exercise in the paper can provide another useful explanation of iDID. More broadly, for methodologists proposing new causal methods, especially those that are historically based on linear models, it may be meaningful to illustrate their new methods under linear models to increase accessibility and accelerate adoption in applied settings.

Acknowledgements

The research of Hyunseung Kang was supported in part by NSF Grant DMS-1811414.

References

Anderson, T. W. and Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* **20,** 46–63.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics.* Princeton University Press.

Dufour, J.-M. (1997). Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica* **65,** 1365–1387.

Guo, Z., Kang, H., Tony Cai, T., and Small, D. S. (2018). Confidence intervals for causal effects with invalid instruments by using two-stage hard thresholding with voting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **80,** 793–815.

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology* **18,** 449–484.

Kang, H., Zhang, A., Cai, T. T., and Small, D. S. (2016). Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. *Journal of the American Statistical Association* **111,** 132–144.

Moreira, M. J. (2009). Tests with correct size when instruments can be arbitrarily weak. *Journal of Econometrics* **152,** 131–140.

Tchetgen Tchetgen, E., Sun, B., and Walter, S. (2021). The GENIUS approach to robust mendelian randomization inference. *Statistical Science* **36,** 443–464.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.

Ye, T., Liu, Z., Sun, B., and Tchetgen Tchetgen, E. (2021). GENIUS-MAWII: For robust mendelian randomization with many weak invalid instruments. *arXiv preprint arXiv:2107.06238* .

Received October 2007. *Revised February* 2008. *Accepted March* 2008.