RESPONSES TO REVIEWER 2

We thank the reviewer for excellent comments. They have caused a major change in the way we think about this problem.

Before getting into our response, let us first explain why this response took so long. There was disagreement amongst the co-authors about the existence of profit-shifting. Further analysis of the original data, following on suggestions the reviewers made, caused some -- but not of all of us -- to change our minds. As a result, this paper has gone through 3 major revisions, each time expanding the original dataset to include more data. As a result of the intensive, ultimately unsuccessful, search for profit-shifting, all of the co-authors became convinced that there is no evidence for profit-shifting in the Chinese data.

Our response proceeds as follows:

- *We address criticisms of the reviewer about our use of the effective tax rate variable (EATR), particularly with respect to concerns about endogeneity. We will present evidence that the Chinese tax system is a flat tax rate with very few exceptions. This minimizes the margin for behavioral responses affecting EATR.*
- *We present the evidence that caused us to conclude that there is no basis for believing that profit-shifting is a problem with Chinese MNCs.*
- *We then respond to the reviewer's remaining comments.*

We would like to have the opportunity to revise our study and resubmit to the journal, this time with the conclusion that there is no evidence of profit-shifting in China. We feel that we have come to this conclusion through a rigorous analysis of the data, spurred by the comments from the reviewers. Given the lack of existing analyses of this important question, we believe our study makes an important contribution to the literature.

ADDRESSING CRITICISM ABOUT THE *EATR* VARIABLE

1. **Comment**: ***"In my opinion, your empirical specification and the conclusions you draw from your results are fundamentally flawed."***

Comment: ***"In this context, it is difficult to interpret the tax elasticities you estimate. And differential investment responses (your outcome variable) to your EATR measure do not help uncover profit shifting, as any profit shifting is already captured in your explanatory variable."***

Response: We apologize for not being clearer about the Chinese tax system and why we believe the backward-looking EATR that we use, while not ideal, is sufficient to identify underlying statutory tax rates. Because this is such an important issue -- perhaps the key concern of the reviewers -- our response is rather detailed. It consists of two parts. First, we give details regarding the Chinese tax system and argue that the backward-looking, effective average tax rate that we use is an unbiased measure of the firm's relevant statutory tax rate, free of behavioral responses. Second, we present new evidence of profit-shifting that causes us to change our conclusion.

Define *EATR* as:

(1) $$EATR = \frac{Taxes\ Paid}{Reported\ Profits}.$$

In a complex corporate tax system like that in the US, *EATR* will be a function of *Reported Profits*. For example, federal corporate income taxes in the US are calculated according to a tiered tax rate schedule. Changes in profits can place the firm in a different tax bracket, affecting the *EATR*. In addition, there is a complex system of deductions, credits, and deferrals all of which can encourage behavioral responses to different tax rates. In this kind of system, the concerns of the reviewers would be completely justified, as profit-shifting would affect *EATRs*.

However, this is not how the Chinese corporate income tax system worked during the time period included in our analysis. Prior to the tax law change of 2008, the Chinese corporate income tax system was regulated by the "Income Tax Law of The People's Republic Of China for Enterprises with Foreign Investment and Foreign Enterprises" ("中华人民共和国外商投资企业和外国企业所得税法"), which was "adopted at the Fourth Session of the Seventh National People's Congress on April 9, 1991, promulgated by Order No. 45 of the President of the People's Republic of China on April 9, 1991 and effective as of July 1, 1991." An English translation of the law can be found here: http://www.law-lib.com/law/law_view.asp?id=7536.

We are just going to excerpt some relevant statutes which we believe will support our contention that the Chinese corporate tax system was 100% flat, with statutory tax rate differences depending on the type of business, geographic location, and other characteristics of the firm (see below, boldfaced added):

Article 4
The taxable income of an enterprise with foreign investment and an establishment or a place set up in China to engage in production or business operations by a foreign enterprise, shall be the amount remaining from its gross income in a tax year after the costs, expenses and losses have been deducted.

Article 5
The income tax on enterprises with foreign investment and the income tax which shall be paid by foreign enterprises on the income of their establishments or places set up in China to engage in production or business operations shall be computed on the taxable income **at the rate of thirty percent**, and local income tax shall be computed on the taxable income **at the rate of three percent**.

Article 7
The income tax on enterprises with foreign investment established in Special Economic Zones, foreign enterprises which have establishments or places in Special Economic Zones engaged in production or business operations, and on enterprises with foreign investment of a production nature in Economic and Technological Development Zones, shall be levied **at the reduced rate of fifteen percent**.

The income tax on enterprises with foreign investment of a production nature established in coastal economic open zones or in the old urban districts of cities where the Special Economic Zones or the Economic and Technological Development Zones are located, shall be levied at **the reduced rate of twenty-four percent**.

The income tax on enterprises with foreign investment in coastal economic open zones, in the old urban districts of cities where the Special Economic Zones or the Economic and Technological Development Zones are located or in other regions defined by the State Council, within the scope of energy, communications, harbour, wharf or other projects encouraged by the State, may be levied **at the reduced rate of fifteen percent**. The specific measures shall be drawn up by the State Council.

Article 8

Any enterprise with foreign investment of a production nature scheduled to operate for a period of not less than ten years shall, from the year beginning to make profit, **be exempted from income tax in the first and second years and allowed a fifty percent reduction in the third to fifth years**. However, the exemption from or reduction of income tax on enterprises with foreign investment engaged in the exploitation of resources such as petroleum, natural gas, rare metals, and precious metals **shall be regulated separately by the State Council**.

More detail is provided in the web link above.

This is the tax law that regulated corporate income taxation in China for FIE's until China passed a new Corporate Income Tax Law in 2007 that went into effect on January 1, 2008 (An and Tan, 2014). Accordingly, we can rewrite Equation (1) as follows:

$$(2)\qquad EATR = \frac{Taxes\ Paid}{Reported\ Profits} = \frac{Reported\ Profits \times Statutory\ Tax\ Rate}{Reported\ Profits}$$
$$= Statutory\ Tax\ Rate.$$

From Equation (2) it can be seen that the ratio of *Taxes Paid* to *Reported Profits* is a measure of the statutory tax rate faced by the firm. While profit-shifting will affect the amount of profits reported by the firm, it will not affect the ratio of *Taxes Paid* to *Reported Profits*. If we are given the opportunity to revise and resubmit our manuscript, we will be sure to emphasize this crucial element of the Chinese corporate tax system.

One way to test this is to regress *ln(PROFITS)* on *EATR*, holding constant other variables that might be correlated both with profits and the firm's eligibility for favorable tax rates. Accordingly, we include all the variables that were used in our original analysis of firms' asset levels. The results are reported below using the same data.

```
Linear regression                               Number of obs     =     20,985
                                                F(15, 12370)      =    1574.86
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6621
                                                Root MSE          =     1.1533

                           (Std. Err. adjusted for 12,371 clusters in firm_id)
------------------------------------------------------------------------------
             |               Robust
   lnPROFITS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        EATR |  -.0223801   .0007501   -29.84   0.000    -.0238504   -.0209098
     lnSALES |   .9869154    .012116    81.46   0.000     .9631661    1.010665
       lnEMP |   .0079702   .0139697     0.57   0.568    -.0194126    .0353529
```

lnGDP	.1809798	.0313986	5.76	0.000	.1194338	.2425259
lnKLRAT	.1637119	.009026	18.14	0.000	.1460195	.1814043
lnSKILL	.1479167	.0201941	7.32	0.000	.1083331	.1875003
lnLABOR	-.1798063	.0549397	-3.27	0.001	-.2874967	-.072116
lnMARK	.3257321	.0516384	6.31	0.000	.2245128	.4269514
lnPOPDEN	-.0445492	.0271664	-1.64	0.101	-.0977996	.0087012
lnINFRAS	-.0824614	.0309452	-2.66	0.008	-.1431187	-.021804
lnFINANCE	-.1772154	.0360081	-4.92	0.000	-.2477968	-.106634
lnINVFOR	-.0322593	.015587	-2.07	0.039	-.0628123	-.0017064
year5	.272274	.0357461	7.62	0.000	.202206	.3423419
year6	.1128163	.0283883	3.97	0.000	.0571708	.1684619
year7	.0744284	.0217891	3.42	0.001	.0317184	.1171384
_cons	-5.574998	.6763577	-8.24	0.000	-6.900765	-4.249232

The coefficient for *EATR* is negative; not zero, as would be expected under a flat tax system; and not positive, as would be expected if Chinese MNCs faced a progressive tax structure. What is responsible for this negative relationship?

One possibility is profit-shifting: Firms facing higher statutory tax rates are more likely to shift their profits overseas. Another possibility is that statutory tax rates are associated with omitted variables that also are correlated with profits. For example, MNCs located in Special Economic Zones are taxed at substantially lower rates (see Article 7 from the Chinese tax law above). If MNCs located in Special Economic Zones also had higher profits, this would induce a negative correlation between profits and *EATR*.

Evidence of the latter is given below, where we drop all variables other than *EATR* and re-estimate the relationship using the same set of observations. Note that the coefficient on *EATR* is almost twice the size in absolute value (-0.042 below versus -0.022 above).

```
Linear regression                                Number of obs     =     20,985
                                                 F(1, 12370)       =    1341.91
                                                 Prob > F          =     0.0000
                                                 R-squared         =     0.0777
                                                 Root MSE          =     1.9047

                           (Std. Err. adjusted for 12,371 clusters in firm_id)
```

lnPROFITS	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
EATR	-.0422478	.0011533	-36.63	0.000	-.0445084	-.0399871
_cons	8.749634	.028331	308.84	0.000	8.694101	8.805167

The results above provide indirect evidence that the negative coefficient on *EATR* could be driven by omitted variable bias. However, they could also be driven by legitimate profit-shifting behavior. This is where the 2-component, finite mixture model approach of EM&W becomes extremely useful. It allows us to go further and directly test if the negative coefficient on *EATR* is being driven by profit-shifting behavior. We do this below in the section that follows this one.

2. <u>**Comment**</u>**:** ***"Your main explanatory variable, the tax measure EATR is as far as I can see a firm-level measure: tax payments relative to pre-tax profits. Using this measure to infer any behavioral responses to tax incentives is wrong, because by definition the measure already captures behavioral responses to tax rates. Firms that do shift profits are going to have a lower individual EATR than firms that do not shift profits. So variation of that variable is endogenous."***

Response: As noted above,

(2) $$EATR = \frac{Taxes\ Paid}{Reported\ Profits} = \frac{Reported\ Profits \times Statutory\ Tax\ Rate}{Reported\ Profits}$$
$$= Statutory\ Tax\ Rate.$$

From Equation (2) it can be seen that the ratio of *Taxes Paid* to *Reported Profits* is a measure of the statutory tax rate faced by the firm. While profit-shifting will affect the amount of profits reported by the firm, it will not affect the ratio of *Taxes Paid* to *Reported Profits*, and thus *EATR* is unaffected by "behavioral responses" and "endogenous choices."

Just so that we are not misunderstood regarding this point, the fact that profits appear on both sides of the regression in Equation (1') does not, per se, induce bias. Running the little Stata .do program below makes that clear.

```
clear
set obs 10000
set seed 13
gen profits = 100*rnormal()
gen taxrate = runiform()
gen taxes = taxrate*profits
// Clearly, taxes/profits = taxrate
drop if profits<0
gen lnprofits = ln(profits)
regress lnprofits taxrate
```

The problem is distinguishing between profit-shifting (i.e., *EATR* increases, *PROFITS* decrease) and omitted variables related both to profits and the statutory tax rate facing the firm. The FMM approach allows us to do this. We discussed this point in the response above and address it in greater detail in the next section.

USING A 2-COMPONENT FMM APPROACH TO DETECT EVIDENCE OF PROFIT-SHIFTING

We now make two major changes from our original paper. First, we put the focus on the responsiveness of profits to tax rates, as this is the most direct connection to profit-shifting. Second, we expand our dataset, including observations through 2009. Our final sample will consist of 57,802 observations, almost three times the sample size of the original. It covers the years 2005-2009, so three years before the tax law change, and two years under the new tax regime. The distribution of observations by year are given below:

	Freq.	Percent	Cum.
2005	9,637	16.67	16.67
2006	11,400	19.72	36.39
2007	13,098	22.66	59.06
2008	11,294	19.54	78.59
2009	12,373	21.41	100.00
Total	57,802	100.00	

The profit-shifting hypothesis requires that the following three predictions be met:

1a) With respect to profits, the FMM should identify two groups of MNCs with different tax elasticities.

1b) Profit-shifters should be more tax elastic than the non-profit-shifters. Profit-shifters are generally expected to be larger firms, as measured by PROFITS, ASSETS, EMPLOYEES, and SALES, because larger firms have the most incentive to profit shift, and are more likely to have overseas connections. Further, we expect the number of profit-shifters to be smaller than the number of non-profit-shifters.

1c) Profit-shifters should evidence a greater tax elasticity effect from the 2008/2009 tax increase than non-profit-shifters.

PREDICTION 1a):
The FMM should identify two groups of MNCs with different tax elasticities

We do find evidence of two groups. The 2-component model has lower AIC and BIC (see below)

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
component1	57,802	-603444.8	-585314.6	18	1170665	1170827

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
component2	57,802	.	-568583.2	37	1137240	1137572

And a LR test rejects the null of a 1-component model

```
Likelihood-ratio test                                 LR chi2(19)  =   33462.81
(Assumption: component1 nested in component2)         Prob > chi2  =     0.0000
```

The two component model produces the following tax elasticities:

```
Class            : 1
Response         : PROFITS
Model            : nbreg, dispersion(constant)

                          (Std. Err. adjusted for 28,044 clusters in FIRM_id)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
PROFITS						
EATR	-.0145755	.0005031	-28.97	0.000	-.0155615	-.0135894

```
Class            : 2
Response         : PROFITS
Model            : nbreg, dispersion(constant)

                          (Std. Err. adjusted for 28,044 clusters in FIRM_id)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
PROFITS						
EATR	-.0106338	.0010798	-9.85	0.000	-.0127501	-.0085174

The tax elasticities of the two groups of firms are statistically different:

```
( 1)  - [PROFITS]1bn.Class#c.EATR + [PROFITS]2.Class#c.EATR = 0

           chi2(  1) =    9.09
         Prob > chi2 =    0.0026
```

CONCLUSION: We find support for Prediction 1a).

PREDICTION 1b):
Profit-shifters should be more tax elastic than non-profit-shifters. Profit-shifters are generally expected to be larger firms, as measured by PROFITS, ASSETS, EMPLOYEES, and SALES, because larger firms have the most incentive to profit shift, and are more likely to have overseas connections. Further, we expect the number of profit-shifters to be smaller than the number of non-profit-shifters.

As reported above, the two component model estimates a larger tax elasticity for “class 1” firms.

```
Class          : 1
Response       : PROFITS
Model          : nbreg, dispersion(constant)

                        (Std. Err. adjusted for 28,044 clusters in FIRM_id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
PROFITS      |
        EATR |  -.0145755   .0005031   -28.97   0.000    -.0155615   -.0135894
```

```
Class          : 2
Response       : PROFITS
Model          : nbreg, dispersion(constant)

                        (Std. Err. adjusted for 28,044 clusters in FIRM_id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
PROFITS      |
        EATR |  -.0106338   .0010798    -9.85   0.000    -.0127501   -.0085174
```

This would suggest that Class 1 firms are the profit-shifters. However, Class 1 firms are more numerous, and have smaller ASSETS, PROFITS, EMP, and SALES (see below).

```
summ ASSETS PROFITS EMP SALES if group1 == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      ASSETS |     50,713    31912.86    127157.3          1   1.22e+07
     PROFITS |     50,716    7659.634    15604.84          2     701859
         EMP |     50,716    304.9162    660.8291          9      52440
       SALES |     50,716    138831.9    500022.8         41   2.93e+07
```

```
summ ASSETS PROFITS EMP SALES if group2 == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      ASSETS |      7,086    332558.1     1108193          5   2.53e+07
     PROFITS |      7,086    152821.2      435020          2   1.47e+07
         EMP |      7,086     1272.56    4737.234          9     198971
       SALES |      7,086     1673586     6537835        890   1.92e+08
```

CONCLUSION: We do not find support for Prediction 1b).
PREDICTION 1c):
Profit-shifters should evidence a greater tax elasticity effect from the 2008/2009 tax increase than non-profit-shifters.

Based on the above, it is unclear which group of firms is the profit-shifters. We created an interaction term (EATRxTD) to pick up the change in tax elasticity after the tax law change. Note that it is insignificant for both groups at the 5 percent level (see below).

```
Class          : 1
Response       : PROFITS
Model          : nbreg, dispersion(constant)

                        (Std. Err. adjusted for 28,044 clusters in FIRM_id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
PROFITS      |
        EATR |  -.0139211   .0006198   -22.46   0.000     -.015136   -.0127063
     EATRxTD |  -.0016905    .000974    -1.74   0.083    -.0035995    .0002185
```

```
Class          : 2
Response       : PROFITS
Model          : nbreg, dispersion(constant)

                        (Std. Err. adjusted for 28,044 clusters in FIRM_id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
PROFITS      |
        EATR |  -.0093657   .0015838    -5.91   0.000    -.0124698   -.0062616
     EATRxTD |  -.0025999    .002149    -1.21   0.226    -.0068118    .0016121
```

Further, there is no statistical difference between the two interaction terms (see below).

```
 ( 1)  - [PROFITS]1bn.Class#c.EATRxTD + [PROFITS]2.Class#c.EATRxTD = 0

           chi2(  1) =    0.12
         Prob > chi2 =    0.7318
```

CONCLUSION: We do not find support for Prediction 1c).

We can also follow EM&W, as our original paper did, and focus on investment behavior. While somewhat indirect, the argument is that profit-shifters, because they are able to shield themselves from taxation, will be relatively tax inelastic compared to non-profit-shifters. Accordingly, we get the following three predictions:

2a) With respect to assets/investments, the FMM should identify two groups of MNCs with different tax elasticities.

2b) Profit-shifters should be less tax elastic than the non-profit-shifters. Profit-shifters are generally expected to be larger firms, as measured by PROFITS, ASSETS, EMPLOYEES, and SALES, because larger firms have the most incentive to profit shift, and are more likely to have overseas connections. Further, we expect the number of profit-shifters to be smaller than the number of non-profit-shifters.

2c) Profit-shifters should evidence a smaller tax elasticity effect from the 2008/2009 tax increase than non-profit-shifters.

PREDICTION 2a):
The FMM should identify two groups of MNCs with different investment/tax elasticities

We do find evidence of two groups. The 2-component model has lower AIC and BIC (see below)

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
component1	57,799	-668213.1	-646851.2	18	1293738	1293900
component2	57,799	.	-622267.5	37	1244609	1244941

And a LR test rejects the null of a 1-component model

```
Likelihood-ratio test                                 LR chi2(19)  =  49167.44
(Assumption: component1 nested in component2)         Prob > chi2  =    0.0000
```

The two component model produces the following investment tax elasticities:

```
Class          : 1
Response       : ASSETS
Model          : nbreg, dispersion(constant)

                          (Std. Err. adjusted for 28,043 clusters in FIRM_id)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ASSETS						
EATR	-.0011314	.0003577	-3.16	0.002	-.0018325	-.0004302

```
Class          : 2
Response       : ASSETS
Model          : nbreg, dispersion(constant)

                          (Std. Err. adjusted for 28,043 clusters in FIRM_id)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ASSETS						
EATR	-.0026553	.0012939	-2.05	0.040	-.0051913	-.0001194

But note that the tax elasticities are NOT statistically different for the two classes:

```
( 1)  - [ASSETS]1bn.Class#c.EATR + [ASSETS]2.Class#c.EATR = 0

           chi2(  1) =     1.16
         Prob > chi2 =    0.2819
```

CONCLUSION: We do not find support for Prediction 2a).

PREDICTION 2b):

Profit-shifters should be LESS tax elastic than the non-profit-shifters. Profit-shifters are generally expected to be larger firms, as measured by PROFITS, ASSETS, EMPLOYEES, and SALES, because larger firms have the most incentive to profit shift, and are more likely to have overseas connections. Further, we expect the number of profit-shifters to be smaller than the number of non-profit-shifters.

As reported above, the two component model estimates a smaller tax elasticity for "Class 1" firms, even though the difference is not statistically significant.

```
Class            : 1
Response         : ASSETS
Model            : nbreg, dispersion(constant)

                          (Std. Err. adjusted for 28,043 clusters in FIRM_id)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ASSETS						
EATR	-.0011314	.0003577	-3.16	0.002	-.0018325	-.0004302

```
Class            : 2
Response         : ASSETS
Model            : nbreg, dispersion(constant)

                          (Std. Err. adjusted for 28,043 clusters in FIRM_id)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ASSETS						
EATR	-.0026553	.0012939	-2.05	0.040	-.0051913	-.0001194

This would suggest that Class 1 firms are the profit-shifters. However, Class 1 firms are more numerous, and have smaller ASSETS, PROFITS, EMP, and SALES (see below).

```
summ ASSETS PROFITS EMP SALES if group1 == 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ASSETS	52,176	24483.72	51515.4	1	1493920
PROFITS	52,176	13302.28	50096.74	2	3020368
EMP	52,176	315.4849	687.7487	9	45306
SALES	52,176	171296	856090	41	6.66e+07

```
. summ ASSETS PROFITS EMP SALES if group2 == 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ASSETS	5,623	479715.6	1257674	5	2.53e+07
PROFITS	5,623	138232.3	476030.9	3	1.47e+07
EMP	5,623	1426.374	5268.03	9	198971
SALES	5,623	1771720	7043604	890	1.92e+08

CONCLUSION: We do not find support for Prediction 2b).

PREDICTION 2c):
Profit-shifters should evidence a smaller tax elasticity effect from the 2008/2009 tax increase than non-profit-shifters.

While we aren't sure which class are the real profit-shifters (if they exist), the interaction term to capture the change in tax elasticity after the tax law change (EATRxTD) is insignificant for both classes (see below).

```
Class            : 1
Response         : ASSETS
Model            : nbreg, dispersion(constant)

                              (Std. Err. adjusted for 28,043 clusters in FIRM_id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ASSETS       |
        EATR |  -.0010082   .0004656    -2.17   0.030    -.0019207   -.0000957
     EATRxTD |  -.0002985   .0006998    -0.43   0.670    -.0016701    .0010731
```

```
Class            : 2
Response         : ASSETS
Model            : nbreg, dispersion(constant)

                              (Std. Err. adjusted for 28,043 clusters in FIRM_id)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ASSETS       |
        EATR |   -.002124   .0020529    -1.03   0.301    -.0061477    .0018996
     EATRxTD |  -.0009903   .0026184    -0.38   0.705    -.0061223    .0041417
```

Further, there is no statistical difference between the two interaction terms (see below).

```
 ( 1)  - [ASSETS]1bn.Class#c.EATRxTD + [ASSETS]2.Class#c.EATRxTD = 0

           chi2(  1) =    0.06
         Prob > chi2 =    0.8092
```

CONCLUSION: We do not find support for Prediction 2c).

FINAL CONCLUSION: The results above have caused us to change our conclusion regarding profit-shifting in China before the law change in 2008. We now conclude that there is no evidence of profit-shifting.

We note that this is not quite as dramatic a change as it might seem, because the original version of the manuscript noted two important caveats (see conclusion of original version). However, the extra analysis motivated by the comments from the reviewers has produced additional results which lead us to this change in conclusion.

RESPONSES TO THE REVIEWER'S REMAINING COMMENTS

3. **Comment**: ***"Given that you closely follow the approach taken in Egger et al., let me lay out some differences with that paper that appear to me to be crucial:"***

Comment: ***"In that paper, tax avoiders (avoidance including not only profit-shifting activities but also other behavioral responses to taxation) are distinguished from non-avoiders because their investments do not react to changes in STATUTORY tax rates (because they are able to avoid taxation)"***

Response: We agree.

4. **Comment**: ***"The EATR measure used there is not the backward-looking measure you use (i.e. what firms actually paid). There, the EATR is constructed using statutory measures such as tax rate, depreciation allowances and other deductions. It is an expected effective tax rate and does not include firm-level choice variables."***

   Response: The reviewer is correct, and our characterization of the EM&W paper's measure of EATR was incorrect. If we are allowed to resubmit a revised version of our manuscript we will make it clear that our analysis uses a backward-looking measure of statutory tax rates. That being said, we believe our analysis above goes far to explaining why we believe our backward-looking measure, while imperfect, still provides a reliable guide to the statutory tax rate facing the individual firms.

5. **Comment**: ***"Of course you cannot exploit variation in statutory tax rates as in Egger et al., given that all the firms in your sample locate in China. But one possible way to proceed would be to exploit the "complex system of preferential tax treatments that cause actual rates to deviate from their statutory values" mentioned on p. 7."***

   Response: We agree that having firm-specific details about preferential tax treatments would be very helpful. Since those are unavailable, we believe our measure of *EATR* is able to substantively recover the relevant statutory tax rate facing the firm.

6. **Comment**: ***"Figure 1 is not relevant."***

   Response: We will drop Figure 1 from a revised version of the manuscript.

7. **Comment**: ***"Equations 1 and 1' are misleading. I would rather argue verbally and include only regression equations that you actually estimate."***

   Response: A revised version of the manuscript will drop Equations (1) and (1').

8. **Comment**: ***"You do not seem to be allowing for unobserved firm-level heterogeneity in your approach (as is done in Egger et al.)"***

   Response: We are concerned about the small number of observations per firm, even in the expanded data set: there are approximately two observations per firm. However, if we are permitted to submit a revised version of our manuscript, we will control for "firm heterogeneity" in the same manner as EM&W as a robustness check.

A FINAL COMMENT TO THE REVIEWER:

In this last section, we wish to emphasize that while our dataset on China is not ideal, and certainly not comparable in quality to what is available to EM&W, we believe we make an important contribution to the study of profit-shifting in China. First, there is a dearth of study on this important subject, driven largely by a lack of Chinese data. The only study on profit-shifting that we are aware of is An and Tan (2014). AT use a difference-in-differences approach to identify the effect of the 2008 tax law change on firms' profits. Their analysis covered the period 2002-2008. Thus, they only have one year of data under the new tax law to identify its effect, and that year coincided with the onset of the global financial crisis (GFC). As they note themselves, because multi-national corporations might be expected to be more impacted by the GFC than domestic firms, there is concern their results may have an identification problem. Second, we believe our paper provides a valuable example for how two-component finite mixture models can be used to provide insight into the existence of profit-shifting.

Finally, while we believe that the flat tax nature of the Chinese tax system goes far in addressing concerns about endogeneity in our tax rate variable, we acknowledge that there may be other, more nuanced ways in which behavioral choices by firms affect our measure of tax rates. If this were the case, it would introduce both measurement error and endogeneity into our tax rate measure. A common approach is to employ the use of instrumental variables in this type of situation. While it is widely known that the cost of reducing bias via IV estimation is that it produces estimates that have greater variance, the size of this cost is often not appreciated. In the Appendix to this response, we have written a Stata program that illustrates this. The DGP is:

$y_i = \beta_0 + \beta_1 x_i + u_i,$

$Corr(x_i, u_i) > 0,$

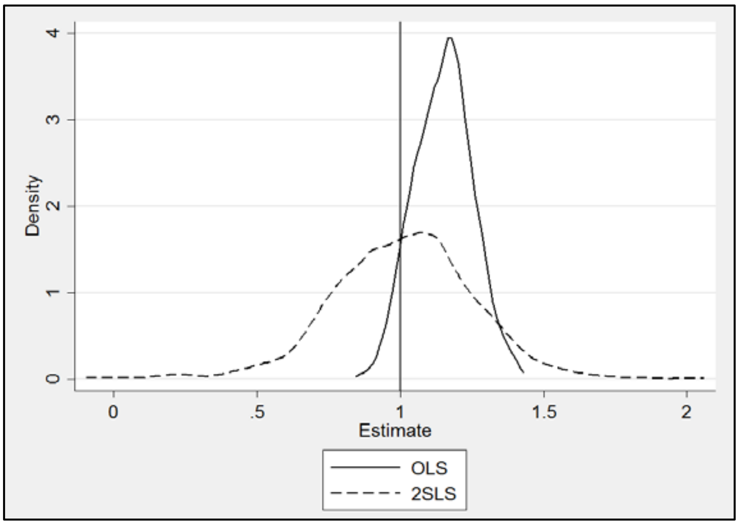
$Corr(z_i, x_i) > 0,$

$Corr(z_i, u_i) = 0.$

The true value of $\beta_1 = 1$. The parameters have been chosen so that endogeneity produces a 15% positive bias to the OLS estimate of β_1. 2SLS eliminates the bias, but at the cost of almost doubling the mean-squared error (see below).

	OLS	2SLS
Bhat	1.1495831	1.0086761
MSE	.03227377	.06095226

This tradeoff between bias and efficiency is illustrated below, with OLS being biased, but also more efficient.



Further, when we test for weak identification (significance level = 5%) using the Cragg-Donald Wald F-statistic, approximately 78% of the time we reject the null of weak identification for “10% maximal size”, and 96% of the time for “15% maximal size”. In other words, the great majority of the time we would conclude that we did not have a weak instrument. Our point is this: every empirical approach is going to have some issues. In light of alternatives and the existing scarce literature on this subject, we believe our approach makes a substantial contribution to the literature.

APPENDIX

```
// This program is the "guts" of the Monte Carlo simulation program.
// You need to first run this program, and then run the second half
// of the program that will make use of this part
set more off
program drop _all
program define IVprog, rclass

syntax, rhoxu(real) rhozx(real) beta0(real) betax(real) ///
        nobs(integer)

// Remove existing variables
drop _all

// Create the data
set obs `nobs'
gen u = rnormal()
gen v = rnormal()
gen w = rnormal()
// The correlation between x and u increases with `rhoxu'
// The reason for dividing through by the sqrt term is to make sure that
// the variance of the x (and z) variables is fixed at one, and does not
// change as we change the rhoxu and rhozx parameters.
gen x = (`rhoxu'*u + v)/sqrt(1+`rhoxu'^2)
// The correlation between z and x increases with `rhozx'
gen z = (`rhozx'*v + w)/sqrt(1+`rhozx'^2)
gen y = `beta0' + `betax'*x + u


// This section calculates OLS estimates
regress y x
return scalar bOLS = _b[x]
test _b[x] = `betax'
return scalar pbOLS = r(p)

// This section calculates 2SLS estimates
ivregress 2sls y (x = z)
return scalar b2SLS = _b[x]
test _b[x] = `betax'
return scalar pb2SLS = r(p)

// This section does the first stage regression and saves the F-stat
regress x z
return scalar Fstat = e(F)

end


drop _all
clear
set more off
graph drop _all
set scheme s2mono
set seed 13
matrix RESULTS = J(2,2,0)
// The local commands below set all the parameters for the experiments.
local rhoxu = 0.15 // Higher values of rhoxu increase the corr(x,u)
local rhozx = 0.5 // Higher values of rhozx increase the corr(z,x)
local beta0 = 1  // This parameters sets the intercept of y/x relationship.
local betax = 1  // This parameter sets the slope coefficient of the x variable.
local nobs = 100
local reps = 1000

// This command runs the simulations.
simulate bOLS = r(bOLS) b2SLS = r(b2SLS) Fstat = r(Fstat), ///
     reps(`reps'): IVprog,  rhoxu(`rhoxu') rhozx(`rhozx') beta0(`beta0') ///
        betax(`betax') nobs(`nobs')

// The next set of commands calculates the average OLS estimates for betax,
// and the MSE of betax

summ bOLS, meanonly
matrix RESULTS[1,1] = r(mean)
```

```
generate msebOLS = (bOLS - `betax')^2
summ msebOLS, meanonly
matrix RESULTS[2,1] = r(mean)

// The set of commands does the same thing for the 2SLS estimates

summ b2SLS, meanonly
matrix RESULTS[1,2] = r(mean)

generate mseb2SLS = (b2SLS - `betax')^2
summ mseb2SLS, meanonly
matrix RESULTS[2,2] = r(mean)

// These commands print out the results
matrix colnames RESULTS = OLS 2SLS
matrix rownames RESULTS = Bhat MSE
matrix list RESULTS

// This command produces a distribution graph of the estimated betax values.
kdensity bOLS , xline(1) name(Combined) ///
    addplot(kdensity b2SLS) legend(label(1 "OLS") ///
        label(2 "2SLS"))

// A popular IV diagnostic is the F-stat from the first-stage regression (Stock and Yogo,
2005)
// The corresponding critical values for the "weak identification test" are:
// 10% maximal IV size = 16.38
// 15% maximal IV size = 8.96

// Accordingly, this last section keeps track of how often the corresponding F-stats
// are greater than 16.38 (="10%" critical value, "check10") and 8.96 (="15%" critical value,
// "check15)

gen check10=(Fstat>16.38)
gen check15=(Fstat>8.96)
summ Fstat check10 check15
```