Discussion of: Marginal Tax Rates and Income: New Time Series Evidence, by K. Mertens

Barbara Rossi
ICREA-Univ. Pompeu Fabra, Barcelona GSE and CREI

ESSIM 2014

Existing Results

- Existing regressions:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

Existing Results

- Existing regressions:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

- where:

Existing Results

- Existing regressions:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

- where:
  - e_t^S = aggregate wage income,

Existing Results

- Existing regressions:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

- where:
  - e_t^S = aggregate wage income,
  - τ_t^s = average marginal tax rate

OLS Regression Results

Table II OLS Regression Results

	All Tax Units Series 1	All Tax Units Series 2	Top 1%	Top 5%	Top 10%	Top 5-1%	Top 10-5%	Btm. 99%	Btm. 90%
				A. *1951-2010*					
Wage Inc.	-0.57^{*}	-0.63^{**}	0.45	0.13	-0.01	-0.15	-0.32	-0.75^{**}	-0.97^{***}
	$(-1.18, 0.03)$	$(-1.19, -0.07)$	$(-0.10, 0.99)$	$(-0.32, 0.59)$	$(-0.43, 0.42)$	$(-0.40, 0.11)$	$(-0.86, 0.22)$	$(-1.33, -0.16)$	$(-1.66, -0.28)$
Total Inc.	-0.33	-0.42	0.58^{*}	0.28	0.14	-0.14	-0.27^{*}	-0.60^{**}	-0.80^{***}
	$(-0.89, 0.23)$	$(-0.95, 0.11)$	$(-0.08, 1.25)$	$(-0.37, 0.94)$	$(-0.44, 0.72)$	$(-0.46, 0.17)$	$(-0.58, 0.04)$	$(-1.08, -0.12)$	$(-1.37, -0.23)$

OLS Regression Results

- Existing regressions:

$$\Delta \ln \left(e_t^s \right) = \varepsilon \Delta \ln \left(1 - \tau_t^s \right) + r_t^s$$

OLS Regression Results

- Existing regressions:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

- For most income groups, estimates are negative, that is tax increases => higher income growth

OLS Regression Results

- Existing regressions:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

- For most income groups, estimates are negative, that is tax increases => higher income growth
- The problem is that there might be endogeneity, and thus the estimate may be biased.

Plan of the Discussion

- I like the main idea: there might be a problem in the existing literature in estimating the relationship between tax rate and income

Plan of the Discussion

- I like the main idea: there might be a problem in the existing literature in estimating the relationship between tax rate and income

- 1. Review of analysis

Plan of the Discussion

- I like the main idea: there might be a problem in the existing literature in estimating the relationship between tax rate and income

- 1. Review of analysis

- 2. Other potentially interesting/useful things to do? Things I am curious about, etc...

Goal of this paper

- Since previous estimates based on OLS might be biased...

Goal of this paper

- Since previous estimates based on OLS might be biased...
- ... the paper provide new estimates of the dynamic effects of marginal tax rate changes on income

Goal of this paper

- Since previous estimates based on OLS might be biased...
- ... the paper provide new estimates of the dynamic effects of marginal tax rate changes on income
- The paper proposes SVAR identification to overcome the bias

Main Contribution

- Estimate a SVAR using exogenous tax shocks.

Main Contribution

- Estimate a SVAR using exogenous tax shocks.
- Let the reduced form VAR be:

$$Y_t = \begin{pmatrix} \ln(e_t) \\ \ln(1-\tau_t) \\ X_t \end{pmatrix}$$

Main Contribution

- Estimate a SVAR using exogenous tax shocks.
- Let the reduced form VAR be:

$$Y_t = \begin{pmatrix} \ln(e_t) \\ \ln(1 - \tau_t) \\ X_t \end{pmatrix}$$

- $$Y_t = A(L)\, Y_{t-1} + u_t$$

Main Contribution

- Estimate a SVAR using exogenous tax shocks.
- Let the reduced form VAR be:

$$Y_t = \begin{pmatrix} \ln(e_t) \\ \ln(1-\tau_t) \\ X_t \end{pmatrix}$$

- $$Y_t = A(L)\,Y_{t-1} + u_t$$

- where X_t = control variables, e_t^S = aggregate wage income, τ_t^s = average marginal tax rate.

Main Contribution

- Estimate a SVAR using exogenous tax shocks.
- Let the reduced form VAR be:

$$Y_t = \begin{pmatrix} \ln(e_t) \\ \ln(1 - \tau_t) \\ X_t \end{pmatrix}$$

- $$Y_t = A(L)\, Y_{t-1} + u_t$$
- where X_t = control variables, e_t^S = aggregate wage income, τ_t^s = average marginal tax rate.
- Let me forget about X_t...

Main Contribution

- Let me forget about X_t...

Main Contribution

- Let me forget about X_t...
- $$Y_t = A(L)\, Y_{t-1} + u_t$$
$$\begin{aligned} Y_t &= \begin{pmatrix} \ln(e_t) \\ \ln(1-\tau_t) \end{pmatrix} \\ u_t &= \begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix}' \end{aligned}$$

Main Contribution

- Let me forget about X_t...
- $$Y_t = A(L)\, Y_{t-1} + u_t$$

$$Y_t = \begin{pmatrix} \ln(e_t) \\ \ln(1-\tau_t) \end{pmatrix}$$
$$u_t = \begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix}'$$

- e_t^S = aggregate wage income, τ_t^s = average marginal tax rate, u_t are reduced-form shocks.

What Do We Care About?

- The structural shocks v_t are such that:

$$u_t = Hv_t$$

What Do We Care About?

- The structural shocks v_t are such that:

$$u_t = Hv_t$$

- where $v_t = (v_t^o, v_t^\tau)'$ have variance identity

What Do We Care About?

- The structural shocks v_t are such that:

$$u_t = Hv_t$$

- where $v_t = (v_t^o, v_t^\tau)'$ have variance identity
- We care about:

What Do We Care About?

- The structural shocks v_t are such that:

$$u_t = Hv_t$$

- where $v_t = (v_t^o, v_t^\tau)'$ have variance identity
- We care about:
  - (i) v_t^τ = the exogenous variation in average marginal tax rates

What Do We Care About?

- The structural shocks v_t are such that:

$$u_t = Hv_t$$

- where $v_t = (v_t^o, v_t^\tau)'$ have variance identity
- We care about:
  - (i) $v_t^\tau =$ the exogenous variation in average marginal tax rates
  - (ii) H_{12} – in fact, e.g. VAR no lags:

$$\begin{pmatrix} \ln(e_t) \\ \ln(1-\tau_t) \end{pmatrix} = \begin{pmatrix} H_{11} & \mathbf{H}_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} v_t^o \\ v_t^\tau \end{pmatrix}$$

Main Contribution

- The structural shocks are estimated using IV (instrumental variables)

Main Contribution

- The structural shocks are estimated using IV (instrumental variables)

- Let m_t = exogenous narrative policy shocks s.t.:

$$\begin{aligned} E\left(v_t^{\tau} m_t\right) &= \alpha \neq 0 \;\; (\textit{relevance}) \\ E\left(v_t^{o} m_t\right) &= 0 \;\; (\textit{exogeneity}) \end{aligned}$$

Technique

- Following Montiel-Olea, Stock and Watson (2012), v_t^τ can be identified by:

Technique

- Following Montiel-Olea, Stock and Watson (2012), v_t^{τ} can be identified by:
- Regress m_t onto u_t and get the **fitted value**: $E\left(m_t u_t\right)' E\left(u_t u_t'\right)^{-1} u_t$

Technique

- Following Montiel-Olea, Stock and Watson (2012), v_t^{τ} can be identified by:
- Regress m_t onto u_t and get the **fitted value**: $E\left(m_t u_t\right)' E\left(u_t u_t'\right)^{-1} u_t$
- $E\left(u_t m_t\right) = E\left(H v_t m_t\right) = HE\left(v_t m_t\right) = HE\begin{pmatrix} v_t^o m_t \\ v_t^{\tau} m_t \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} 0 \\ \alpha \end{pmatrix} = \begin{pmatrix} H_{12} \\ H_{22} \end{pmatrix}\alpha = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix}\alpha$

Technique

- Following Montiel-Olea, Stock and Watson (2012), v_t^τ can be identified by:
- Regress m_t onto u_t and get the **fitted value**: $E\left(m_t u_t\right)' E\left(u_t u_t'\right)^{-1} u_t$
- $E\left(u_t m_t\right) = E\left(H v_t m_t\right) = HE\left(v_t m_t\right) = HE\begin{pmatrix} v_t^o m_t \\ v_t^\tau m_t \end{pmatrix} =$ $\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} 0 \\ \alpha \end{pmatrix} = \begin{pmatrix} H_{12} \\ H_{22} \end{pmatrix}\alpha = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix}\alpha$
- ($H_{22} = 1$ is normalization)

Technique

- Following Montiel-Olea, Stock and Watson (2012), v_t^{τ} can be identified by:
- Regress m_t onto u_t and get the **fitted value**: $E\left(m_t u_t\right)' E\left(u_t u_t'\right)^{-1} u_t$
- $E\left(u_t m_t\right) = E\left(H v_t m_t\right) = HE\left(v_t m_t\right) = HE\begin{pmatrix} v_t^o m_t \\ v_t^{\tau} m_t \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} 0 \\ \alpha \end{pmatrix} = \begin{pmatrix} H_{12} \\ H_{22} \end{pmatrix}\alpha = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix}\alpha$
- ($H_{22} = 1$ is normalization)
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH'$

Technique

- Following Montiel-Olea, Stock and Watson (2012), v_t^τ can be identified by:
- Regress m_t onto u_t and get the **fitted value**: $E\left(m_t u_t\right)' E\left(u_t u_t'\right)^{-1} u_t$
- $E\left(u_t m_t\right) = E\left(H v_t m_t\right) = HE\left(v_t m_t\right) = HE\begin{pmatrix} v_t^o m_t \\ v_t^\tau m_t \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} 0 \\ \alpha \end{pmatrix} = \begin{pmatrix} H_{12} \\ H_{22} \end{pmatrix}\alpha = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix}\alpha$
- ($H_{22} = 1$ is normalization)
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH'$
- **Fitted value**: $E\left(m_t u_t\right)' E\left(u_t u_t'\right)^{-1} u_t = \underbrace{\alpha \begin{pmatrix} H_{12} \\ 1 \end{pmatrix}' H'^{-1}}_{=[0,1]} \underbrace{H^{-1} u_t}_{v_t} = \alpha\left[0,1\right] v_t = \alpha v_t^\tau$

Technique II

- $E\left(u_t m_t\right) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix} \alpha$

Technique II

- $E\left(u_t m_t\right) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix} \alpha$
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH'$

Technique II

- $E\left(u_t m_t\right) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix} \alpha$
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH^{'}$
- Fitted value: αv_t^τ

Technique II

- $E\left(u_t m_t\right) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix}\alpha$
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH^{'}$
- Fitted value: αv_t^τ
- Thus:

Technique II

- $E(u_t m_t) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix} \alpha$
- $E(u_t u_t') = E(H v_t v_t' H') = HH'$
- Fitted value: αv_t^τ
- Thus:
  - (i) $E(u_t^\tau m_t) = \alpha$

Technique II

- $E\left(u_t m_t\right) = E\left(\left(\begin{array}{c} u_t^e \\ u_t^\tau \end{array}\right) m_t\right) = \left(\begin{array}{c} H_{12} \\ 1 \end{array}\right) \alpha$
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH'$
- Fitted value: αv_t^τ
- Thus:
  - (i) $E\left(u_t^\tau m_t\right) = \alpha$
  - (ii) the regression m_t onto u_t gives fitted values $=\alpha\ v_t^\tau$ from which v_t^τ can be estimated

Technique II

- $E\left(u_t m_t\right) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix} \alpha$
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH'$
- Fitted value: αv_t^τ
- Thus:
  - (i) $E\left(u_t^\tau m_t\right) = \alpha$
  - (ii) the regression m_t onto u_t gives fitted values $=\alpha\ v_t^\tau$ from which v_t^τ can be estimated
  - (iii) $E\left(u_t^e m_t\right) = H_{12}\alpha$ gives an estimate of H_{12}

Technique II

- $E\left(u_t m_t\right) = E\left(\begin{pmatrix} u_t^e \\ u_t^\tau \end{pmatrix} m_t\right) = \begin{pmatrix} H_{12} \\ 1 \end{pmatrix} \alpha$
- $E\left(u_t u_t'\right) = E\left(H v_t v_t' H'\right) = HH'$
- Fitted value: αv_t^τ
- Thus:
  - (i) $E\left(u_t^\tau m_t\right) = \alpha$
  - (ii) the regression m_t onto u_t gives fitted values $=\alpha\ v_t^\tau$ from which v_t^τ can be estimated
  - (iii) $E\left(u_t^e m_t\right) = H_{12}\alpha$ gives an estimate of H_{12}
  - (iv) estimate impulse responses by tracing the effect of v_t^τ on Y_t.

Question I. About the Estimation

- The paper estimates instead a 3-stage regression. I am not sure why the the paper proceeds in this way. This seems to complicate the analysis/estimation.

Question I. About the Estimation

- The paper estimates instead a 3-stage regression. I am not sure why the the paper proceeds in this way. This seems to complicate the analysis/estimation.

- One of the complications following the three stage procedure is that impulse response bands are quite convoluted to calculate, and the author has to rely on Goncalves and Kilian's bootstrap (2004), whereas Montiel-Olea, Stock and Watson's procedure provides a simple formula...

Question I. About the Estimation

- The paper estimates instead a 3-stage regression. I am not sure why the the paper proceeds in this way. This seems to complicate the analysis/estimation.

- One of the complications following the three stage procedure is that impulse response bands are quite convoluted to calculate, and the author has to rely on Goncalves and Kilian's bootstrap (2004), whereas Montiel-Olea, Stock and Watson's procedure provides a simple formula...

- Bottom line: why not comparing the results of alternative procedures? Are there advantages relative to the simpler procedure?

Three-stage Regression Estimated in This Paper

- Regress u_t^e on u_t^τ using m_t as instruments, and get the residuals

Three-stage Regression Estimated in This Paper

- Regress u_t^e on u_t^τ using m_t as instruments, and get the residuals
- Regress u_t^τ onto u_t^e using the residuals in the previous step as instruments, and get the residual, which is v_t^τ

Three-stage Regression Estimated in This Paper

- Regress u_t^e on u_t^τ using m_t as instruments, and get the residuals
- Regress u_t^τ onto u_t^e using the residuals in the previous step as instruments, and get the residual, which is v_t^τ
- Regress the variables you are interested in onto $v_t^\tau, v_{t-1}^\tau, \ldots$ to get IRFs

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.
- A good instrument needs to be **exogenous**:

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.
- A good instrument needs to be **exogenous**:
  - $\Rightarrow E\left(v_t^o m_t\right) = 0$ $\left(\textit{exogeneity}\right)$

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.
- A good instrument needs to be **exogenous**:
  - $\Rightarrow E\left(v_t^o m_t\right) = 0$ $\left(exogeneity\right)$
- A good instrument needs to be **relevant**:

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.
- A good instrument needs to be **exogenous**:
  - $\Rightarrow E\left(v_t^o m_t\right) = 0$ $\left(exogeneity\right)$
- A good instrument needs to be **relevant**:
  - $\Rightarrow E\left(v_t^\tau m_t\right) = \alpha \neq 0$ $\left(relevance\right)$

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.
- A good instrument needs to be **exogenous**:
  - $\Rightarrow E\left(v_t^o m_t\right) = 0 \; (exogeneity)$
- A good instrument needs to be **relevant**:
  - $\Rightarrow E\left(v_t^\tau m_t\right) = \alpha \neq 0 \; (relevance)$
  - I did not see evidence on the relevance condition: it would be important to add evidence on this (i.e. first-stage F-stat of regressing u_t^τ on m_t) – if relevance does not hold, instrument is weak and significance may be spurious

Question II. About the Instruments

- 1. The methodology relies on m_t being a good instrument.
- A good instrument needs to be **exogenous**:
  - $\Rightarrow E\left(v_t^o m_t\right) = 0$ $(exogeneity)$
- A good instrument needs to be **relevant**:
  - $\Rightarrow E\left(v_t^\tau m_t\right) = \alpha \neq 0$ $(relevance)$
  - I did not see evidence on the relevance condition: it would be important to add evidence on this (i.e. first-stage F-stat of regressing u_t^τ on m_t) – if relevance does not hold, instrument is weak and significance may be spurious
  - The paper reports first-stage F-tests on elasticities estimated from $\Delta \ln\left(e_t^s\right) = \varepsilon \Delta \ln\left(1 - \tau_t^s\right) + r_t^s$ using v_t^τ as instrument – more later...

Question III. Model Specification: Nonlinearities

- Here we are imposing that one additional contemporaneous unanticipated shock in taxes v_t^{τ} has a constant (linear) effect H_{12} on the log of aggregate income $ln\left(e_t\right)$.

Question III. Model Specification: Nonlinearities

- Here we are imposing that one additional contemporaneous unanticipated shock in taxes v_t^{τ} has a constant (linear) effect H_{12} on the log of aggregate income $ln\left(e_t\right)$.

- However, the responses may be nonlinear (threshold models)

Question III. Model Specification: Nonlinearities

- Here we are imposing that one additional contemporaneous unanticipated shock in taxes v_t^τ has a constant (linear) effect H_{12} on the log of aggregate income $ln\left(e_t\right)$.

- However, the responses may be nonlinear (threshold models)

- Test whether there is empirical evidence of non-linearities?

Question IV: Heterogeneity Analysis

- Heterogeneity:

Question IV: Heterogeneity Analysis

- Heterogeneity:
  1. the paper also estimates responses to MTR changes at different income levels:

$$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

using v_t^τ as instrument => estimates are now positive!

Question IV: Heterogeneity Analysis

- Heterogeneity:
  1. the paper also estimates responses to MTR changes at different income levels:

     $$\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$$

     using v_t^τ as instrument => estimates are now positive!

  2. the paper also looks at dynamic responses of different income percentiles to the shock:

     $$\Delta \ln \left(e_t^s\right) = \gamma_0 v_t^\tau + \gamma_1 v_{t-1}^\tau + \ldots$$

Question IV: Heterogeneity Questions

1. Why not estimating a SVAR for each income group separately (controlling for that income group's lagged income, otherwise the estimated shock is not necessarily exogenous to the group's income – and confidence bands cannot be calculated...)

Question IV: Heterogeneity Questions

1. Why not estimating a SVAR for each income group separately (controlling for that income group's lagged income, otherwise the estimated shock is not necessarily exogenous to the group's income – and confidence bands cannot be calculated...)

2. Are they different across income levels? Confidence bands?

Question IV: Heterogeneity Questions

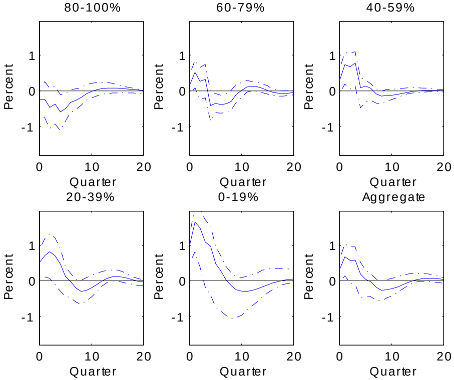
1. Why not estimating a SVAR for each income group separately (controlling for that income group's lagged income, otherwise the estimated shock is not necessarily exogenous to the group's income – and confidence bands cannot be calculated...)

2. Are they different across income levels? Confidence bands?

3. The paper reports first-stage F-tests on elasticities estimated from $\Delta \ln \left(e_t^s\right) = \varepsilon \Delta \ln \left(1 - \tau_t^s\right) + r_t^s$ using v_t^τ as instrument – but v_t^τ is estimated in the aggregate VAR – has the first stage F-statistics been corrected for that?

- 4. Heterogeneity. Anderson, Inoue and Rossi (2012) show that the response of consumption is very different across income groups.

IRF of Consumption to Gov. Spending Shock by Income Group



Question V. Additional Checks?

- The individuals in the top 1% might have different characteristics (age, ...) – is it possible that the response of the top 1% be different depending on their characteristics? Would that be interesting to investigate?

Question V. Additional Checks?

- The individuals in the top 1% might have different characteristics (age, ...) – is it possible that the response of the top 1% be different depending on their characteristics? Would that be interesting to investigate?
- I wonder if it would be useful to support the claims in this paper with micro evidence based on survey data (if available)

Conclusions

- Very interesting idea

Conclusions

- Very interesting idea
- I would like to better understand the advantages of the method

Conclusions

- Very interesting idea
- I would like to better understand the advantages of the method
- Hope the comments are useful