Lecture 5: Marginal Treatment Effects and the relationship between evaluation parameters

Costas Meghir

January 2009

Marginal Treatment Effects - Motivation

- We have seen that different evaluation parameters are an average over parts of the distribution of impacts.

Marginal Treatment Effects - Motivation

- We have seen that different evaluation parameters are an average over parts of the distribution of impacts.
  - The ATE averages over the entire distribution

Marginal Treatment Effects - Motivation

- We have seen that different evaluation parameters are an average over parts of the distribution of impacts.
  - The ATE averages over the entire distribution
  - The ATT averages over the distribution of impacts for those who are *somehow* allocated into treatment

Marginal Treatment Effects - Motivation

- We have seen that different evaluation parameters are an average over parts of the distribution of impacts.
  - The ATE averages over the entire distribution
  - The ATT averages over the distribution of impacts for those who are *somehow* allocated into treatment
  - LATE averages over the distribution of impacts for those who *switch* into treatment as a result of a reform or more precisely, as a result of a change of the value of some instrument affecting decisions to participate.

- Thus they all represent an aggregation over *different margins*

Marginal Treatment Effects - Motivation

- Thus they all represent an aggregation over *different margins*
- As such they are not comparable and they are difficult to interpret from the perspective of general

Marginal Treatment Effects - Motivation

- Thus they all represent an aggregation over *different margins*
- As such they are not comparable and they are difficult to interpret from the perspective of general
- As a unifying parameter Heckman and Vytlacil (2005) defined the *MARGINAL TREATMENT EFFECT*

Marginal Treatment Effects - Motivation

- Thus they all represent an aggregation over *different margins*
- As such they are not comparable and they are difficult to interpret from the perspective of general
- As a unifying parameter Heckman and Vytlacil (2005) defined the *MARGINAL TREATMENT EFFECT*
- This is the effect of a treatment on the marginal individual entering treatment

Marginal Treatment Effects - Motivation

- Thus they all represent an aggregation over *different margins*
- As such they are not comparable and they are difficult to interpret from the perspective of general
- As a unifying parameter Heckman and Vytlacil (2005) defined the *MARGINAL TREATMENT EFFECT*
- This is the effect of a treatment on the marginal individual entering treatment
- The marginal treatment effect will provide an interpretation of several evaluation parameters

Marginal Treatment Effects - Motivation

- Thus they all represent an aggregation over *different margins*
- As such they are not comparable and they are difficult to interpret from the perspective of general
- As a unifying parameter Heckman and Vytlacil (2005) defined the *MARGINAL TREATMENT EFFECT*
- This is the effect of a treatment on the marginal individual entering treatment
- The marginal treatment effect will provide an interpretation of several evaluation parameters
- They will provide a bridge between structural an treatment effect parameters and allow us to understand the way they are related.

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.
- The rule allocating to treatment may be written as

$$T = 1\left(v_i \leq Z_i'\gamma\right)$$

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.
- The rule allocating to treatment may be written as

$$T = 1\left(v_i \leq Z_i'\gamma\right)$$

- For a particular value of $Z_i'\gamma$ the marginal individual is the one with $v_i = Z_i'\gamma$

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.
- The rule allocating to treatment may be written as

$$T = 1\left(v_i \leq Z_i'\gamma\right)$$

- For a particular value of $Z_i'\gamma$ the marginal individual is the one with $v_i = Z_i'\gamma$
- Now consider the effect of treatment for the ith individual $\beta_i = Y_i^1 - Y_i^0$

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.
- The rule allocating to treatment may be written as

$$T = 1\left(v_i \leq Z_i'\gamma\right)$$

- For a particular value of $Z_i'\gamma$ the marginal individual is the one with $v_i = Z_i'\gamma$
- Now consider the effect of treatment for the ith individual $\beta_i = Y_i^1 - Y_i^0$
- The *marginal treatment* effect is defined by

$$MTE(Z_i'\gamma) = E(\beta_i | v_i = Z_i'\gamma)$$

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.
- The rule allocating to treatment may be written as

$$T = 1\left(v_i \leq Z_i'\gamma\right)$$

- For a particular value of $Z_i'\gamma$ the marginal individual is the one with $v_i = Z_i'\gamma$
- Now consider the effect of treatment for the ith individual $\beta_i = Y_i^1 - Y_i^0$
- The *marginal treatment* effect is defined by

$$MTE(Z_i'\gamma) = E(\beta_i | v_i = Z_i'\gamma)$$

- Thus the MTE is the average impact for the marginal individual receiving treatment among those with value of the index equal to $Z_i'\gamma$

Marginal Treatment Effects - a Definition

- Consider a discrete treatment T.
- The rule allocating to treatment may be written as

$$T = 1\left(v_i \leq Z_i'\gamma\right)$$

- For a particular value of $Z_i'\gamma$ the marginal individual is the one with $v_i = Z_i'\gamma$
- Now consider the effect of treatment for the ith individual $\beta_i = Y_i^1 - Y_i^0$
- The *marginal treatment* effect is defined by

$$MTE(Z_i'\gamma) = E(\beta_i | v_i = Z_i'\gamma)$$

- Thus the MTE is the average impact for the marginal individual receiving treatment among those with value of the index equal to $Z_i'\gamma$
- It turns out that all parameters we have looked at can be written as weighted averages of this parameter

Marginal Treatment Effects - a simple justification.

- Now Suppose we think of a very simple model of College attendance choice.

Marginal Treatment Effects - a simple justification.

- Now Suppose we think of a very simple model of College attendance choice.
- The Cost of attending is $C_i = W_0 + Z_i' \gamma + u_{si}$ where W_0 is the opportunity cost in lost earnings and $Z_i' \gamma + u_{si}$ represents the direct costs

Marginal Treatment Effects - a simple justification.

- Now Suppose we think of a very simple model of College attendance choice.
- The Cost of attending is $C_i = W_0 + Z_i'\gamma + u_{si}$ where W_0 is the opportunity cost in lost earnings and $Z_i'\gamma + u_{si}$ represents the direct costs
- The benefits in the simplest form are $L(W_i^1 - W_i^0)$ where L represents a lifecycle factor and depends on the discount rate

Marginal Treatment Effects - a simple justification.

- Now Suppose we think of a very simple model of College attendance choice.
- The Cost of attending is $C_i = W_0 + Z_i'\gamma + u_{si}$ where W_0 is the opportunity cost in lost earnings and $Z_i'\gamma + u_{si}$ represents the direct costs
- The benefits in the simplest form are $L(W_i^1 - W_i^0)$ where L represents a lifecycle factor and depends on the discount rate
- An individual will go to College if $L(W_i^1 - W_i^0) \geq W_0 + Z_i'\gamma + u_{si}$

Marginal Treatment Effects - a simple justification.

- Now Suppose we think of a very simple model of College attendance choice.
- The Cost of attending is $C_i = W_0 + Z_i'\gamma + u_{si}$ where W_0 is the opportunity cost in lost earnings and $Z_i'\gamma + u_{si}$ represents the direct costs
- The benefits in the simplest form are $L(W_i^1 - W_i^0)$ where L represents a lifecycle factor and depends on the discount rate
- An individual will go to College if $L(W_i^1 - W_i^0) \geq W_0 + Z_i'\gamma + u_{si}$
- The marginal individual satisfies the condition $u_{si} = L(W_i^1 - W_i^0) - W_0 - Z_i'\gamma$

Marginal Treatment Effects - a simple justification.

- Now Suppose we think of a very simple model of College attendance choice.
- The Cost of attending is $C_i = W_0 + Z_i'\gamma + u_{si}$ where W_0 is the opportunity cost in lost earnings and $Z_i'\gamma + u_{si}$ represents the direct costs
- The benefits in the simplest form are $L(W_i^1 - W_i^0)$ where L represents a lifecycle factor and depends on the discount rate
- An individual will go to College if $L(W_i^1 - W_i^0) \geq W_0 + Z_i'\gamma + u_{si}$
- The marginal individual satisfies the condition $u_{si} = L(W_i^1 - W_i^0) - W_0 - Z_i'\gamma$
- This shows how the allocation to treatment will depend on the returns and why conditioning on u_s will give us the treatment effect for the marginal individual at a given $Z_{i.}$

Marginal Treatment Effects

- It is convenient to rewrite the treatment model as

$$T = 1\left(u_{si} \leq F(Z_i'\gamma)\right)$$

where u_{si} is now uniform [0,1].

Marginal Treatment Effects

- It is convenient to rewrite the treatment model as

$$T = 1\left(u_{si} \leq F(Z_i'\gamma)\right)$$

  where u_{si} is now uniform [0,1].
- This can be done by defining $U_s \equiv F(V)$ where $F(\cdot)$ is the distribution function of V (lower case denotes a specific realisation of the random variable in upper case). Thus $F(Z'\gamma) = \Pr(T = 1|Z) = P(Z)$.

Marginal Treatment Effects

- It is convenient to rewrite the treatment model as

$$T = 1\left(u_{si} \leq F(Z_i'\gamma)\right)$$

  where u_{si} is now uniform [0,1].
- This can be done by defining $U_s \equiv F(V)$ where $F(\cdot)$ is the distribution function of V (lower case denotes a specific realisation of the random variable in upper case). Thus $F(Z'\gamma) = \Pr(T = 1|Z) = P(Z)$.
- Define

$$\begin{aligned} Y_0 &= \gamma_0'X + U_0 \\ Y_1 &= \gamma_1'X + U_1 \end{aligned}$$

Marginal Treatment Effects

- It is convenient to rewrite the treatment model as

$$T = 1\left(u_{si} \leq F(Z_i'\gamma)\right)$$

  where u_{si} is now uniform [0,1].
- This can be done by defining $U_s \equiv F(V)$ where $F(\cdot)$ is the distribution function of V (lower case denotes a specific realisation of the random variable in upper case). Thus $F(Z'\gamma) = \Pr(T = 1|Z) = P(Z)$.
- Define

$$Y_0 = \gamma_0' X + U_0$$
$$Y_1 = \gamma_1' X + U_1$$

- Now consider the observed outcome

$$Y = \gamma_0' X + T(\gamma_1 - \gamma_0)' X + U_0 + T(U_1 - U_0)$$

Marginal Treatment Effects

- We are now going to think of the probability of assignment to treatment $P(Z)$ as an instrument.

Marginal Treatment Effects

- We are now going to think of the probability of assignment to treatment $P(Z)$ as an instrument.
- Assume that
$$(U_0, U_1, U_s) \perp\!\!\!\perp P(Z) | X$$

Marginal Treatment Effects

- We are now going to think of the probability of assignment to treatment $P(Z)$ as an instrument.
- Assume that

$$(U_0, U_1, U_s) \perp\!\!\!\perp P(Z)|X$$

- We will also make the assumption that *given* X $P(Z)$ has continuous support in the open interval (0,1): this means that the excluded variables Z vary sufficiently for *any* fixed value of X to make the treatment assignment probability vary anywhere between 0 and 1.

- Now take the expected value of the outcome given the instrument $P(Z) = p$ and X.

$$E(Y|X, P(Z) = p) =$$

$$\gamma_0' X + p(\gamma_1 - \gamma_0)' X + E\left[T(U_1 - U_0)|X, P(Z) = p\right]$$

Marginal Treatment Effects - Local Instrumental Variables

- Now take the expected value of the outcome given the instrument $P(Z) = p$ and X.

$$E(Y|X, P(Z) = p) =$$

$$\gamma_0' X + p(\gamma_1 - \gamma_0)' X + E\left[T(U_1 - U_0)|X, P(Z) = p\right]$$

- Now note that $T = 1$ over the interval for $u_s = [0, p]$ and zero for higher values of u_s and u_s is uniform

$$E\left[T(U_1 - U_0)|P(Z) = p, X\right]$$

$$= \int_{-\infty}^{+\infty} \int_0^p (U_1 - U_0) f((U_1 - U_0)|U_s = u_s) du_s d(U_1 - U_0)$$

Marginal Treatment Effects - Local Instrumental Variables

- We can now write the marginal treatment effect as

$$\Delta^{MTE}(p) = \frac{\partial E(Y|X,P(Z)=p)}{\partial p} =$$

$$(\gamma_1 - \gamma_0)'X + \int_{-\infty}^{+\infty} (U_1 - U_0) f((U_1 - U_0)|U_s = p) d(U_1 - U_0)$$

$$(\gamma_1 - \gamma_0)'X + E(U_1 - U_0|U_s = p)$$

Marginal Treatment Effects - Local Instrumental Variables

- We can now write the marginal treatment effect as

$$\Delta^{MTE}(p) = \frac{\partial E(Y|X,P(Z)=p)}{\partial p} =$$

$$(\gamma_1 - \gamma_0)'X + \int_{-\infty}^{+\infty}(U_1 - U_0)f((U_1 - U_0)|U_s = p)d(U_1 - U_0)$$

$$(\gamma_1 - \gamma_0)'X + E(U_1 - U_0|U_s = p)$$

- $E(U_1 - U_0|U_s = p)$ is the average unobserved gain of treatment for those whose unobserved characteristics make them indifferent between treatment or not at $P(Z) = p$.

- The above process suggests an estimation procedure:

Marginal Treatment Effects - Local Instrumental Variables

- The above process suggests an estimation procedure:
- Estimate the nonparametric regression of the outcome variable Y on X and on $P(Z)$.This can be achieved by fitting Y on polynomials of X and $P(Z)$.

Marginal Treatment Effects - Local Instrumental Variables

- The above process suggests an estimation procedure:
- Estimate the nonparametric regression of the outcome variable Y on X and on $P(Z)$.This can be achieved by fitting Y on polynomials of X and $P(Z)$.
- Differentiate the result with respect to $P(Z)$

Marginal Treatment Effects - Local Instrumental Variables

- The above process suggests an estimation procedure:
- Estimate the nonparametric regression of the outcome variable Y on X and on $P(Z)$.This can be achieved by fitting Y on polynomials of X and $P(Z)$.
- Differentiate the result with respect to $P(Z)$
- If $P(Z)$ indeed varies from $(0, 1)$, i.e. has full support, then it will be possible to estimate the marginal treatment effect

Marginal Treatment Effects and evaluation parameters

- We can now define all parameters of interest as a function of the MTE

Marginal Treatment Effects and evaluation parameters

- We can now define all parameters of interest as a function of the MTE
- Averaging over all marginal individuals we obtain

$$
\begin{aligned}
ATE(X) &= \int_{-\infty}^{\infty} \Delta^{MTE}(p)dp \\
& (\gamma_1 - \gamma_0)'X + \int_{-\infty}^{\infty} E(U_1 - U_0|u_s)du_s \\
&= (\gamma_1 - \gamma_0)'X
\end{aligned}
$$

Marginal Treatment Effects and evaluation parameters

- Now consider LATE. Here for any given X $P(Z)$ takes two values, say $P(Z) = b(X)$ or $P(Z) = a(X)$ with say $a > b$. So lets think of $a(X) - b(X)$ as the policy induced change in the treatment probability for someone with characteristics X.

Marginal Treatment Effects and evaluation parameters

- Now consider LATE. Here for any given X $P(Z)$ takes two values, say $P(Z) = b(X)$ or $P(Z) = a(X)$ with say $a > b$. So lets think of $a(X) - b(X)$ as the policy induced change in the treatment probability for someone with characteristics X.
- All those with unobserved propensity to be assigned to treatment such that $b(X) \leq u_s \leq a(X)$ will now switch into treatment under this policy

Marginal Treatment Effects and evaluation parameters

- Now consider LATE. Here for any given X $P(Z)$ takes two values, say $P(Z) = b(X)$ or $P(Z) = a(X)$ with say $a > b$. So lets think of $a(X) - b(X)$ as the policy induced change in the treatment probability for someone with characteristics X.
- All those with unobserved propensity to be assigned to treatment such that $b(X) \leq u_s \leq a(X)$ will now switch into treatment under this policy
- Hence LATE can be written as

$$LATE(X) = \frac{1}{a(X)-b(X)} \int_{b(X)}^{a(X)} \Delta^{MTE}(p)dp$$

$$(\gamma_1 - \gamma_0)'X + \frac{1}{a(X)-b(X)} \int_{b(X)}^{a(X)} E(U_1 - U_0|u_s)du_s$$

Marginal Treatment Effects and evaluation parameters

- Now consider LATE. Here for any given X $P(Z)$ takes two values, say $P(Z) = b(X)$ or $P(Z) = a(X)$ with say $a > b$. So lets think of $a(X) - b(X)$ as the policy induced change in the treatment probability for someone with characteristics X.
- All those with unobserved propensity to be assigned to treatment such that $b(X) \leq u_s \leq a(X)$ will now switch into treatment under this policy
- Hence LATE can be written as

$$LATE(X) = \frac{1}{a(X)-b(X)} \int_{b(X)}^{a(X)} \Delta^{MTE}(p)dp$$

$$(\gamma_1 - \gamma_0)'X + \frac{1}{a(X)-b(X)} \int_{b(X)}^{a(X)} E(U_1 - U_0|u_s)du_s$$

- The value of LATE will depend on the interval over which we integrate, i.e. it will depend on which margin the specific policy tend to shift into treatment.

Marginal Treatment Effects and evaluation parameters

- Heckman and Vytlacil (2005) show that every estimator can be written as a weighted average of the MTE.

Marginal Treatment Effects and evaluation parameters

- Heckman and Vytlacil (2005) show that every estimator can be written as a weighted average of the MTE.
- For example LATE is the average MTE with weights $\frac{1}{a(X)-b(X)}$ in the range $(a(X), b(X))$ and zero everywhere else

Marginal Treatment Effects and evaluation parameters

- Heckman and Vytlacil (2005) show that every estimator can be written as a weighted average of the MTE.
- For example LATE is the average MTE with weights $\frac{1}{a(X)-b(X)}$ in the range $(a(X), b(X))$ and zero everywhere else
- Consider now the parameter Treated on the treated. This can be written as

$$\beta^{TT}(X) = \int_{-\infty}^{+\infty} \Delta^{MTE}(p) \left[\frac{\Pr(P(Z|X) > p)}{E(P(Z|X))} \right] dp$$

Marginal Treatment Effects and evaluation parameters

- Heckman and Vytlacil (2005) show that every estimator can be written as a weighted average of the MTE.
- For example LATE is the average MTE with weights $\frac{1}{a(X)-b(X)}$ in the range $(a(X), b(X))$ and zero everywhere else
- Consider now the parameter Treated on the treated. This can be written as

$$\beta^{TT}(X) = \int_{-\infty}^{+\infty} \Delta^{MTE}(p) \left[\frac{\Pr(P(Z|X) > p)}{E(P(Z|X)} \right] dp$$

- Thus the higher the probability of participating the larger the weight on the overall effect

Marginal Treatment Effects and evaluation parameters

- Now consider the difference between the IV and the OLS estimator. In a model with homogeneous impacts if the treatment is positively correlated with the error term then IV will be biased upwards [exercise to show]:

Marginal Treatment Effects and evaluation parameters

- Now consider the difference between the IV and the OLS estimator. In a model with homogeneous impacts if the treatment is positively correlated with the error term then IV will be biased upwards [exercise to show]:
- One key result in Heckman and Vytlacil is that in models with *heterogeneous* impacts, where selection into treatment depends on the impact this result is no longer true:

Marginal Treatment Effects and evaluation parameters

- Now consider the difference between the IV and the OLS estimator. In a model with homogeneous impacts if the treatment is positively correlated with the error term then IV will be biased upwards [exercise to show]:
- One key result in Heckman and Vytlacil is that in models with *heterogeneous* impacts, where selection into treatment depends on the impact this result is no longer true:
- They show this by writing IV and OLS as different weighted averages of the MTE

Marginal Treatment Effects and evaluation parameters

- We drop X for simplicity of notation

Marginal Treatment Effects and evaluation parameters

- We drop X for simplicity of notation
- For IV the weights are:
  $w^{IV}(u_s) = \left[E(P(Z)|P(Z) > u_s) - E\left(P(Z)\right)\right] \frac{E(P(Z))}{Var(P(Z))}$

Marginal Treatment Effects and evaluation parameters

- We drop X for simplicity of notation
- For IV the weights are:
  $w^{IV}(u_s) = \left[E(P(Z)|P(Z) > u_s) - E\left(P(Z)\right)\right] \frac{E(P(Z))}{Var(P(Z))}$
- For OLS the weights are $w^{OLS}(u_s) = 1 + \frac{E(U_1|U_S=u_s)h_1 - E(U_0|U_S=u_s)h_0}{\Delta^{MTE}(u_s)}$ if $\Delta^{MTE}(u_s) \neq 0$ and *zero*otherwise.

Marginal Treatment Effects and evaluation parameters

- We drop X for simplicity of notation
- For IV the weights are:
  $w^{IV}(u_s) = [E(P(Z)|P(Z) > u_s) - E(P(Z))] \frac{E(P(Z))}{Var(P(Z))}$
- For OLS the weights are $w^{OLS}(u_s) = 1 + \frac{E(U_1|U_S = u_s)h_1 - E(U_0|U_S = u_s)h_0}{\Delta^{MTE}(u_s)}$ if $\Delta^{MTE}(u_s) \neq 0$ and *zero*otherwise.
- In the above $h_1 = E(P(Z)|P(Z) > u_s)/E(P(Z))$ and $h_0 = E(P(Z)|P(Z) < u_s)/E(P(Z))$

Marginal Treatment Effects and evaluation parameters

- We drop X for simplicity of notation
- For IV the weights are:
  $w^{IV}(u_s) = [E(P(Z)|P(Z) > u_s) - E(P(Z))] \frac{E(P(Z))}{Var(P(Z))}$
- For OLS the weights are $w^{OLS}(u_s) = 1 + \frac{E(U_1|U_S=u_s)h_1 - E(U_0|U_S=u_s)h_0}{\Delta^{MTE}(u_s)}$ if $\Delta^{MTE}(u_s) \neq 0$ and *zero*otherwise.
- In the above $h_1 = E(P(Z)|P(Z) > u_s)/E(P(Z))$ and $h_0 = E(P(Z)|P(Z) < u_s)/E(P(Z))$
- Under monotonicity all IV weights are positive.

Marginal Treatment Effects and evaluation parameters

- We drop X for simplicity of notation
- For IV the weights are:
  $w^{IV}(u_s) = [E(P(Z)|P(Z) > u_s) - E(P(Z))] \frac{E(P(Z))}{Var(P(Z))}$
- For OLS the weights are $w^{OLS}(u_s) = 1 + \frac{E(U_1|U_S=u_s)h_1 - E(U_0|U_S=u_s)h_0}{\Delta^{MTE}(u_s)}$ if $\Delta^{MTE}(u_s) \neq 0$ and *zero*otherwise.
- In the above $h_1 = E(P(Z)|P(Z) > u_s)/E(P(Z))$ and $h_0 = E(P(Z)|P(Z) < u_s)/E(P(Z))$
- Under monotonicity all IV weights are positive.
- However, OLS weights may be negative and there is no specific relationship to the IV weighs that will make one estimator larger or smaller than the other.

The Returns to College (Carneiro, Heckman and Vytlacil)

- Object of study: Estimate the returns to College and analyse the heterogeneity of these returns.

The Returns to College (Carneiro, Heckman and Vytlacil)

- Object of study: Estimate the returns to College and analyse the heterogeneity of these returns.
- Data NLSY 1979

The Returns to College (Carneiro, Heckman and Vytlacil)

- Object of study: Estimate the returns to College and analyse the heterogeneity of these returns.
- Data NLSY 1979
- Outcome variable log wages

The Returns to College (Carneiro, Heckman and Vytlacil)

- Object of study: Estimate the returns to College and analyse the heterogeneity of these returns.
- Data NLSY 1979
- Outcome variable log wages
- Conditioning variables (X): Years of experience, Cognitive ability (AFQT), Maternal Education, Cohort dummies, log average Earnings in SMSA, and average unemployment rate in State.

The Returns to College (Carneiro, Heckman and Vytlacil)

- Object of study: Estimate the returns to College and analyse the heterogeneity of these returns.
- Data NLSY 1979
- Outcome variable log wages
- Conditioning variables (X): Years of experience, Cognitive ability (AFQT), Maternal Education, Cohort dummies, log average Earnings in SMSA, and average unemployment rate in State.
- Instruments (Z): Presence of a four year public College in SMSA at age 14, log average earnings in the SMSA when 17 (opportunity cost),average unemployment rate in State.

The Returns to College (Carneiro, Heckman and Vytlacil)

Table 2
Sample Statistics

	$S = 0$ ($N = 717$)	$S = 1$ ($N = 903$)
Log Hourly Wage	2.4029	2.7406
	(0.5568)	(0.5493)
Years of Experience	10.1838	7.5162
	(4.2233)	(3.9804)
Corrected AFQT	-0.3580	0.5563
	(0.8806)	(0.7650)
Mother's Years of Schooling	11.4895	12.8992
	(2.0288)	(2.2115)
SMSA Log Earnings in 1994	10.2707	10.3277
	(0.1618)	(0.1738)
State Unemployment in 1994	5.7793	5.9292
(in %)	(1.2431)	(1.2851)
Presence of a College at 14	0.4616	0.5825
	(0.4988)	(0.4934)
SMSA Log Earnings at 17	10.2793	10.2760
	[illegible]	[illegible]

The Returns to College (Carneiro, Heckman and Vytlacil)

- Estimate a logit model for College participation on cohort dummies and on polynomial terms of the instruments

The Returns to College (Carneiro, Heckman and Vytlacil)

- Estimate a logit model for College participation on cohort dummies and on polynomial terms of the instruments
- The Probability of College attendance then is

$$P(Z) \equiv \Pr(T = 1|Z) = \frac{1}{1 + \exp(-Z'\beta)}$$

The Returns to College (Carneiro, Heckman and Vytlacil)

- Estimate a logit model for College participation on cohort dummies and on polynomial terms of the instruments
- The Probability of College attendance then is

$$P(Z) \equiv \Pr(T = 1|Z) = \frac{1}{1 + \exp(-Z'\beta)}$$

- The average derivatives are then simply

$$\frac{1}{N} \sum_{Sample} \left[\frac{\partial Pr(T = 1|Z)}{\partial AFQT} \right] = \frac{1}{N} \sum_{Sample} \left[P(Z)(1 - P(Z)) \frac{\partial Z'\beta}{\partial AFQT} \right]$$

Table 3
Average Derivatives for College Decision Model

Corrected AFQT	0.2238
	(0.0279)
Mother's Years of Schooling	0.0422
	(0.0119)
Presence of a College at 14	0.0933
	(0.0231)
SMSA Log Earnings at 17	-0.1543
	(0.0761)

The Returns to College (Carneiro, Heckman and Vytlacil)

$$
\begin{aligned}
\ln w &= a + \beta(3.5 \times T) + \gamma' X + u \\
\ln w &= a + \beta(X)(3.5 \times T) + \gamma' X + u
\end{aligned}
$$

Table 4

OLS and IV Estimates of the Return to One Year of College

	Return does not vary with X						Return varies with X	
	OLS	IV					OLS	IV
		Distance	Earnings	Unemployment	All	P		P
β	0.0389	0.1896	0.2431	0.0787	0.1865	0.1379	0.0502	0.1751
	(0.0087)	(0.0960)	(0.1230)	(0.1301)	(0.0573)	(0.0470)	(0.0119)	(0.0661)
$\partial\beta/\partial$AFQT							0.0249	0.0855
							(0.0148)	(0.0385)
F - Statistic (first stage)		2.79	1.90	1.31	2.63	2.23		2.23
p-value		0.01	0.07	0.25	0.00	0.00		0.00

Notice: a. How the results vary by Instrumental Variable and b. How IV is larger than OLS.

The Returns to College (Carneiro, Heckman and Vytlacil)

Consider now the various different evaluation parameters

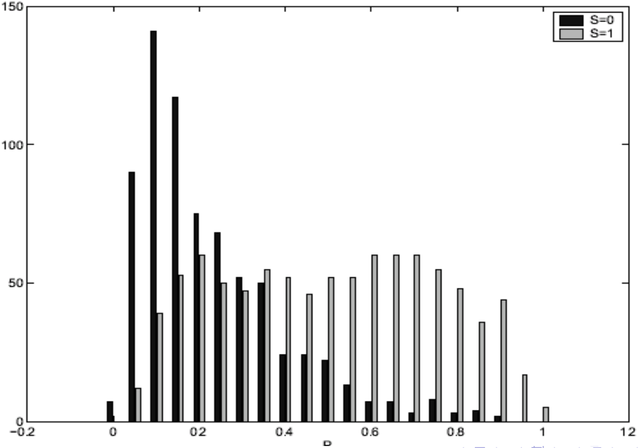
Table 5
Estimates of Various Returns to One Year of College
(Semi-Parametric Model)

	$0.0541 < P < 0.9662$
Average Treatment Effect	0.1832
	(0.0855)
Treatment on the Treated	0.2165
	(0.0978)
Treatment on the Untreated	0.1672
	(0.0875)
Average Marginal Treatment Effect	0.1793
	(0.1114)
Policy Relevant Treatment Effect	0.2013
(Construction of Colleges)	(0.1079)
Ordinary Least Squares	0.0502
	(0.0119)
Instrumental Variables	0.1751
	(0.0661)

The Returns to College (Carneiro, Heckman and Vytlacil)

Support and identification: How does the probability of going to College differ between those who go to College and those who do not?

Figure 1: Density of P Given $S = 0$ and $S = 1$
(Estimated Probability of Enrolling in College)



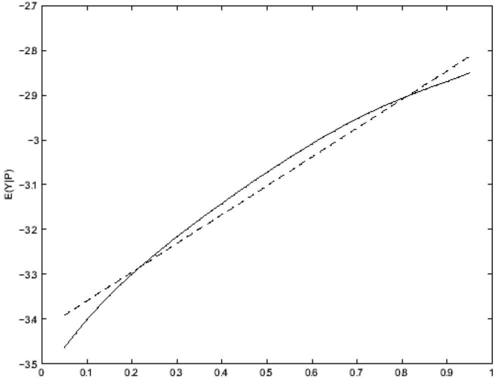
The Returns to College (Carneiro, Heckman and Vytlacil)

Are returns Heterogeneous? - direct evidence

$$E(Y|X, P(Z) = p) =$$

$$\gamma_0' X + p(\gamma_1 - \gamma_0)' X + E\left[T(U_1 - U_0)|X, P(Z) = p\right]$$

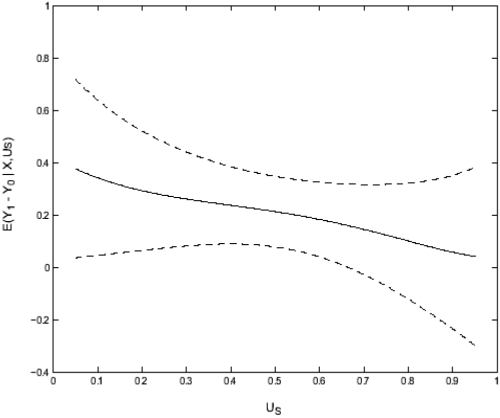
Figure 2: $E(Y|X,P)$ as a Function of P for Average X



The Returns to College (Carneiro, Heckman and Vytlacil)

The Marginal Treatment Effect by unobserved cost of College

Figure 3: $E(Y_1 - Y_0|X, U_S)$ Estimated Using Locally Quadratic Regression (Averaged Over X)



The Returns to College (Carneiro, Heckman and Vytlacil)

The weights implied by IV and OLS

Figure 10: OLS and IV Weights for $E(Y_1 - Y_0|X, U_S)$
(Averaged Over X)