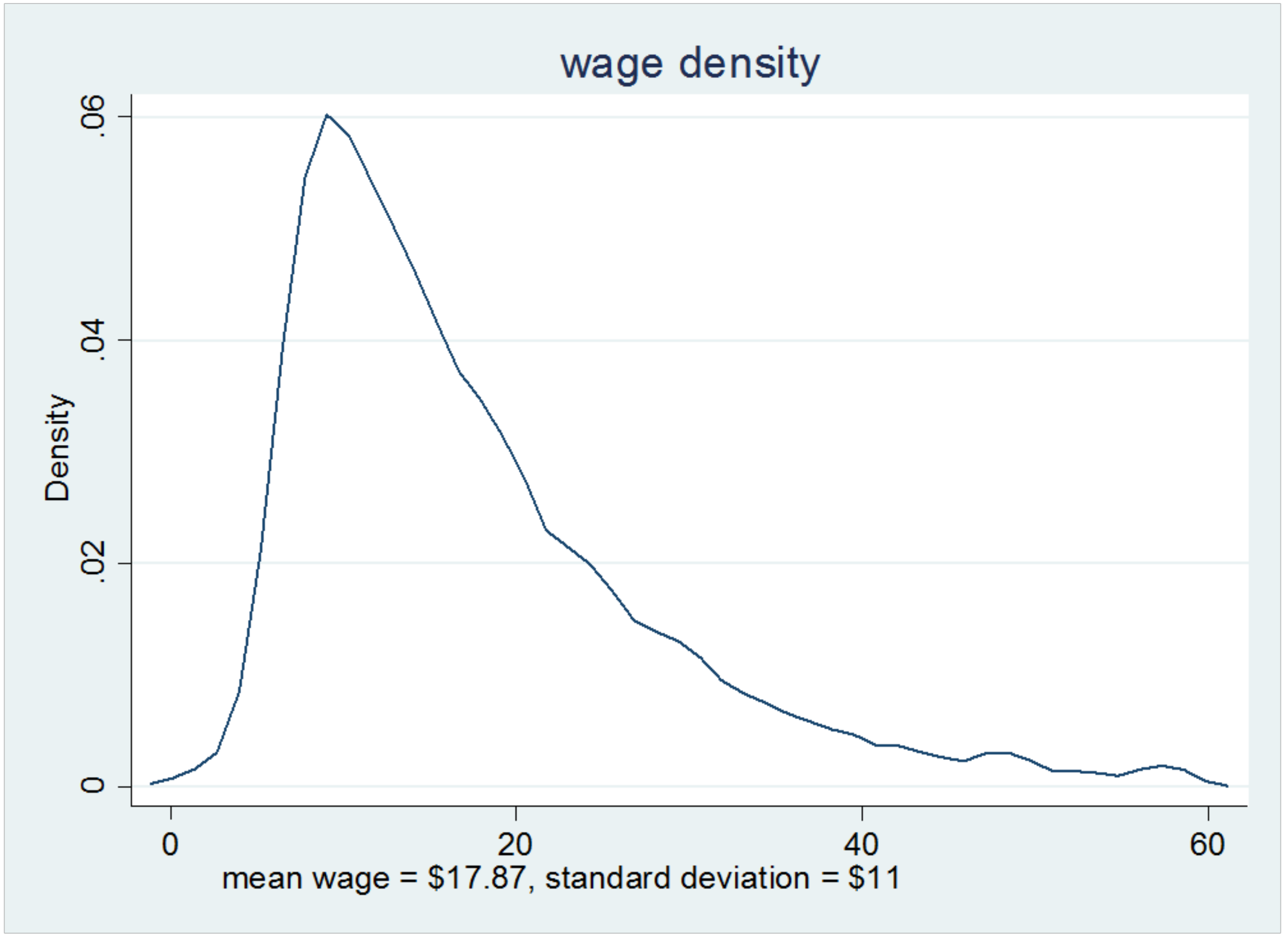
Forecasting

- A *forecast* is a guess about an unknown.
- An *economic forecast* is a forecast about an economic variable, event, outcome, or duration.

Let’s Make a Forecast

- Suppose we take a random household in the United States.
- Let’s forecast the wage (hourly) of the head of household.
- What is your forecast?
- Will your forecast be correct? Why?

Wage Density



Wage Distribution

Wage Forecast

- Wages have a distribution in the population.
- It is impossible to correctly forecast an individual's wage.
- If we forecast "the wage will be \$17.87" it is close to impossible that a given person's wage will be exactly \$17.87.
- The most correct and accurate forecast is the entire distribution (or density).

Forecast Distribution

Percentiles

1%	$ 4.37
5%	$ 6.27
10%	$ 7.50
25%	$10.00
50%	$14.76
75%	$22.45
90%	$32.30
95%	$40.60
99%	$57.70

Forecast Distribution

- Suppose we are trying to forecast an economic variable *y*
- For example, a random person's wage.
- *y* has a distribution *F(y)* which is mathematically defined as

$$F(u) = P(y \leq u)$$

- Visually, we represent distributions through their density functions $f(y) = \frac{d}{dy} F(y)$

Forecast Distribution

- A complete forecast for *y* is its distribution *F(y)* or density *f(y)*.
- Either *F(y)* or *f(y)* summarizes all that is known and unknown about the potential values for *y*.
- We call *F* the *forecast* or *predictive* distribution.
- Can you forecast a person's wage?
  - We cannot know with certainty the wage
  - We know (or can estimate) the distribution:
    - The range and likelihood of possible wages.

What might this matter?

- Suppose your company underwrites unemployment insurance which pay a person's wage y if they become unemployed.
- Suppose a random person loses their job.
- What is the cost to the company?
- We cannot know with certainty, but we may know the *distribution* of the potential costs.

Point Forecast

- In many applications, users want a single number.
- A *point forecast* $\hat{y}$ is our best guess for the unknown object.
- A point forecast for a random person’s wage could be the mean \$17.87 or the median \$14.76

Point Forecast

- A point forecast $\hat{y}$ is a function of the predictive distribution F
- It can be viewed as a summary of F.
- Which function should be used?
- What is our best guess for y based on F?
- It turns out that the answer depends on our *loss function* – how we measure the costs due to potential forecast error.

Forecast Error

- If we forecast a random variable y with a forecast f we say that the *forecast error* is
$$e = y - \hat{y}$$
- The forecast error is the difference between the actual and the forecast.
- For example, if we forecasted that an individual's wage would be \$18, but it turns out that it is \$24, then the error is 24-18=6. If their wage was actually \$14, then the error would be 14-18=-4.

Forecast error

- So long as the variable y is random (not perfectly forecastable) then there will always be forecast error.
- This cannot be avoided.
- However, errors are costly.
- A user can assign costs to a forecast error.
- We call this the loss function

$$L(e) = Loss\ Function$$

Loss Functions

- Common mathematical choices
  - Quadratic Loss

$$L(e) = e^2$$

  - Absolute Loss

$$L(e) = |e|$$

- Both are symmetric
  - Treat positive and negative forecast errors symmetrically
- Quadratic loss penalizes large errors much more than small errors.
- Asymmetric Loss Functions also possible.

Many Loss Functions are (Approximately) Quadratic

- Consider a monopolist selling a product Q at a price P, with linear demand and zero cost.
- The monopolist sets price P and then sells Q.
- The demand equation is $Q=2a-P$
- The profit function is $\pi(P)=2aP-P^2$
- The optimal price is $P^*=a$, optimal profit $\pi^*=\pi(P^*)=a^2$.
- Let $\hat{a}$ be a forecast of a with error $e=a-\hat{a}$.
- The monopolist sets $P=\hat{a}$
- The Loss is $L=\pi^*-\pi(\hat{a})= a^2 -2a\hat{a}+\hat{a}^2=(a-\hat{a})^2=e^2$
- This is quadratic loss.

Risk

- The *Risk* of a forecast is its expected loss.
- Mathematically,

$$R(\hat{y}) = \mathrm{E}\,L(e) = \mathrm{E}\,L(y - \hat{y})$$

- For quadratic loss

$$R(\hat{y}) = \mathrm{E}\left(y - \hat{y}\right)^2$$

- For absolute loss

$$R(\hat{y}) = \mathrm{E}\left|y - \hat{y}\right|$$

Optimal Point Forecast

- The optimal (best) point forecast is the function $\hat{y}$ of the predictive distribution F which minimizes the risk (minimizes the expected loss).
- In the quadratic case

$$\begin{aligned} R(\hat{y}) &= \mathrm{E}\left(y-\hat{y}\right)^2 \\ &= \mathrm{E}\, y^2 - 2\hat{y}\,\mathrm{E}\, y + \hat{y}^2 \end{aligned}$$

which is a quadratic in $\hat{y}$

Optimal Point Forecast – Quadratic Loss

- The $\hat{y}$ which minimizes the Risk is found by differentiation

$$0 = -2\,\mathrm{E}\,y + 2\hat{y}$$

- Which has the solution

$$\hat{y} = \mathrm{E}\,y$$

- The optimal point forecast is the mean of the predictive distribution

Optimal Forecast Under Quadratic Loss is the Mean

- The optimal point forecast under quadratic loss is the mean.
- For example, to forecast the wage of a random person, our optimal point forecast is $17.87

Optimal Prediction Under Absolute Loss is the Median

- The risk of a forecast is

$$R(\hat{y}) = \mathrm{E}\left|y - \hat{y}\right|$$

- This is minimized by the median
- The optimal forecast under absolute loss is

$$\hat{y} = Median(y)$$

- For example, to forecast the wage of a random person, the optimal point forecast is \$14.76

Choice of Loss Function

- We have learned that the optimal point forecast depends upon the loss function
- The mean Ey minimizes the expected squared error
- The median minimizes the expected absolute error.
- Other loss functions lead to different solutions.
- In most cases, we do not have an explicit loss function. So we take the simplest approach and use the mean, which is equivalent to squared loss.
- However, in a real-world application, you might be able to articulate the explicit loss due to forecasting error. In this case, it would be best to use the loss function explicitly, leading to specialized estimators and forecasts.

Interval Forecast

- We have said that a complete forecast for the unknown wage y is its density f, but for simplicity users often want a point forecast $\hat{y}$
- An intermediate solution is to report a forecast interval $C=[\hat{y}_L, \hat{y}_U]$.
- A forecast interval is similar to a confidence interval in statistics.
- The goal is for the unknown wage y to lie in the forecast interval with a pre-specified probability.

Interval Forecast

- Thus an *x*% forecast interval C satisfies

$$P\left(y \in C\right) = x$$

- Common choices for *x* include
  - $x = .90$ (90%)
  - $x = .80$ (80%)
  - $x = .50$ (50%)
- 50% intervals have the simple property that they contain the unknown *y* with even odds.

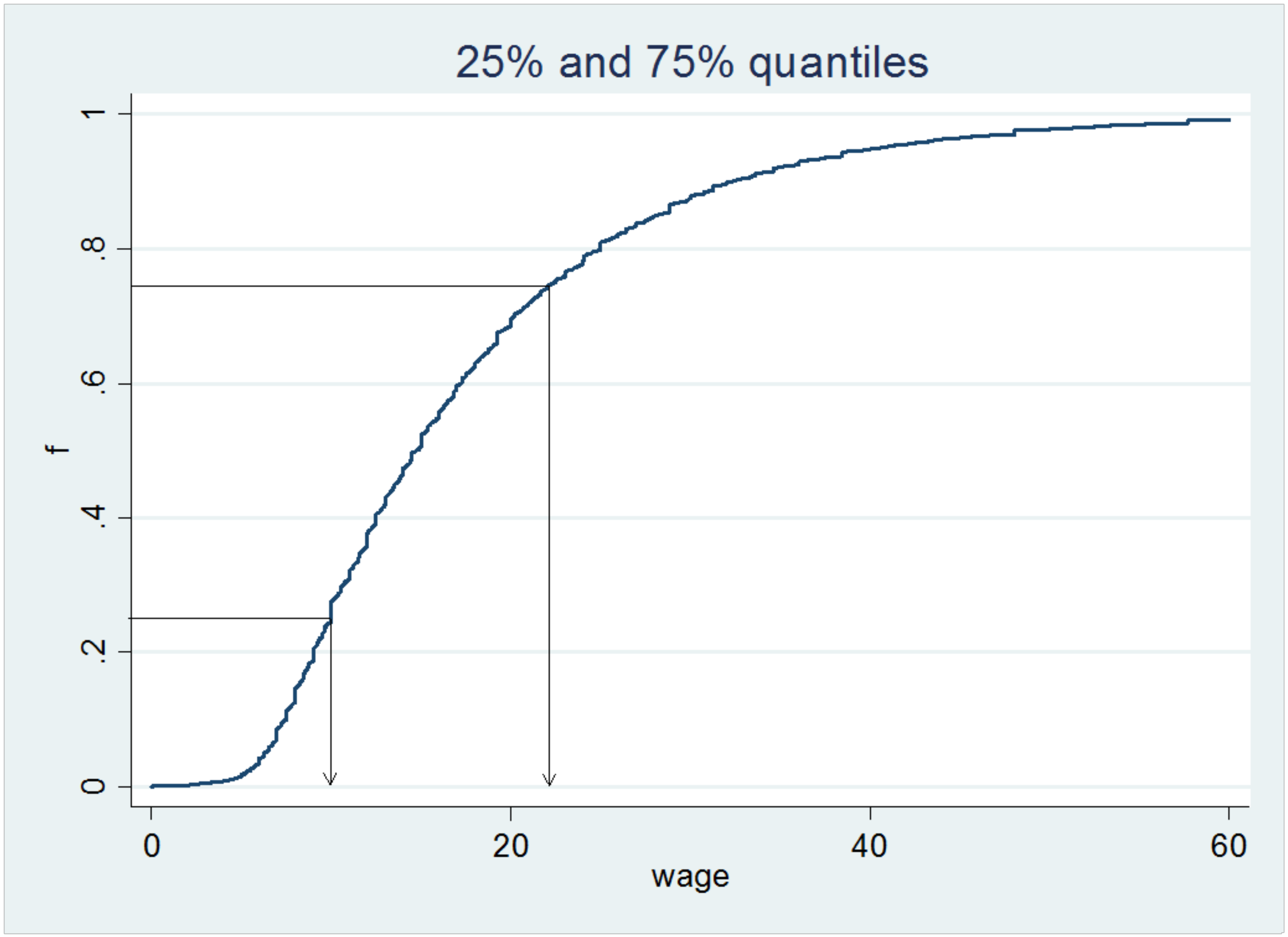
Quantiles

- The endpoints of $C=[\hat{y}_L, \hat{y}_U]$ are quantiles of the distribution F of y.
- Definition: The α'th quantile of y is the number q_α which satisfies

$$\alpha = F(q_\alpha)$$

- They are found by inverting the distribution function
- For a x% interval, you need the $x/2$ and 1-$x/2$ quantiles
- For example, the 25% and 75% quantiles for a 50% forecast interval.

Quantiles of Wage Distribution



Normal Rule

- If the variable *y* is normally distributed $N(\mu,\sigma^2)$
  - The point forecast is μ
  - The forecast intervals are $[\mu-\sigma z_{\alpha/2}, \mu+\sigma z_{\alpha/2}]$ where $z_{\alpha/2}$ are quantiles from the normal distribution table.
  - For example, for a 90% interval, $z_{.05}=1.645$, or for a 50% interval, $z_{.25}=0.675$
- All you need to know is the standard deviation
- But economic data are often far from normal, so this rule may be inaccurate.

Use Forecast Intervals!

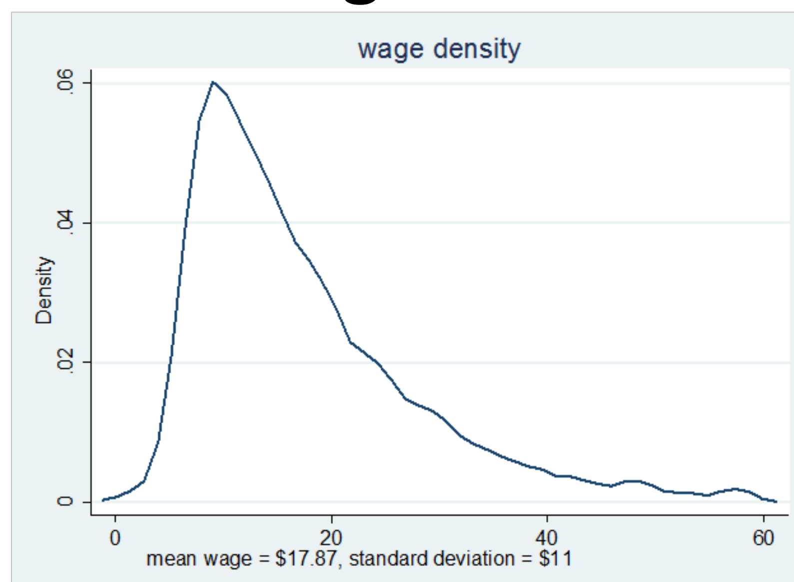
- Forecast intervals are simple, yet not widely used.
- A point forecast by itself does not communicate the uncertainty in the forecast
- A forecast interval is easier to interpret than the entire distribution

Summary – Unconditional Forecasts

- A complete forecast of a random variable y is the distribution F or density f of the variable.
- A point forecast is a single number $\hat{y}$ to summarize the distribution.
- The optimal choice depends on the loss function.
- When loss is quadratic, the optimal point forecast is the mean.
- Forecast intervals are quantiles of the forecast distribution, and convey useful information about the uncertain in y.

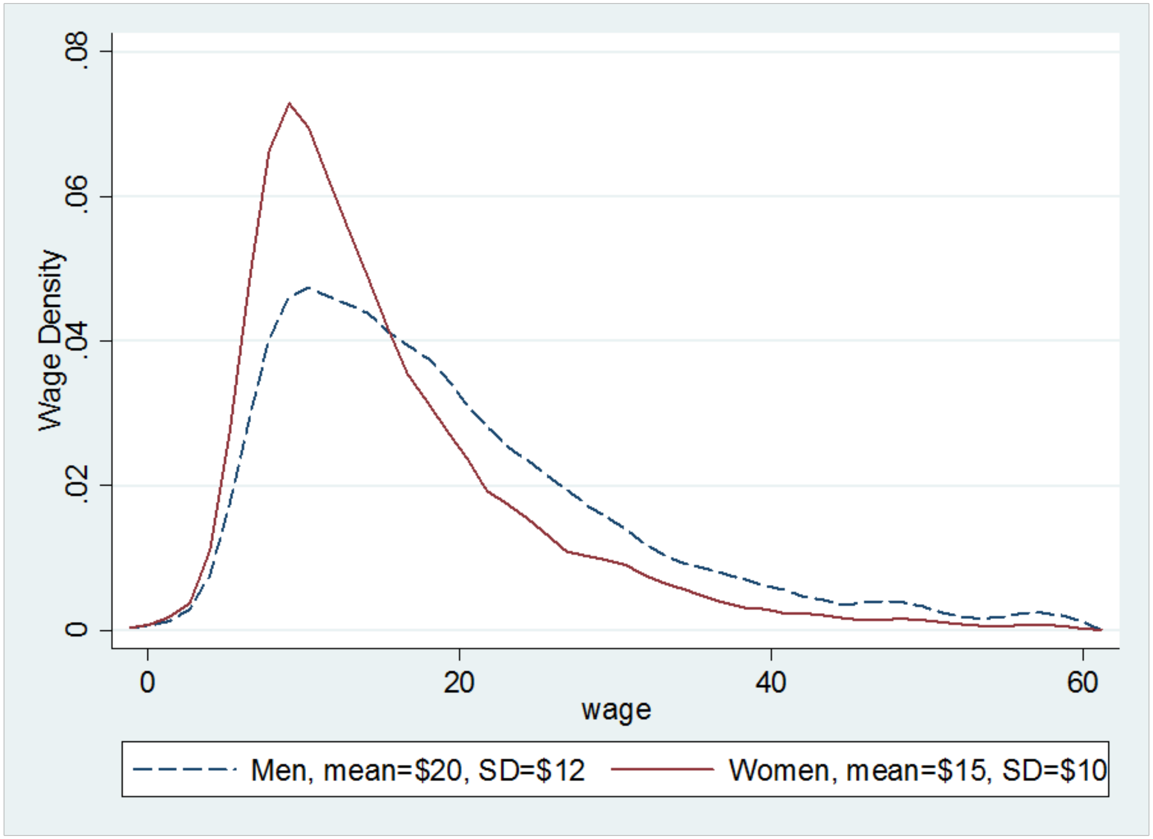
Conditional Forecast

- We had considered forecasting the wage of a random person.
- The distribution is quite diffuse as it includes all wage earners. We know nothing about the person being forecast.

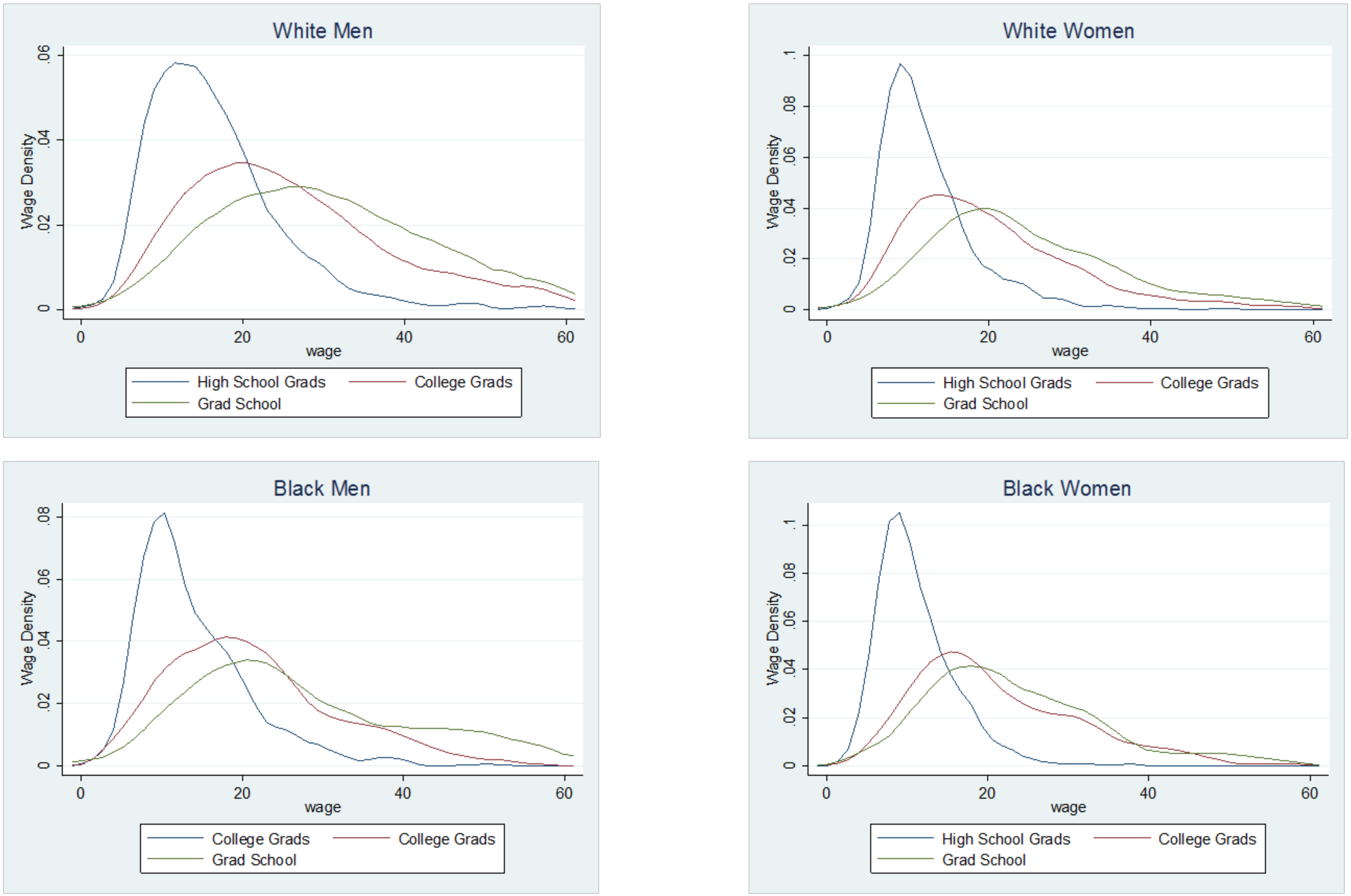


Conditional Forecast

- Now suppose we know that the person is a man (or a woman).
- The information improves the forecast.



Conditional on Sex, Race, Education



Conditional Forecasts (Means)

		Men	**Women**
White	High School	$17	$13
	College	$27	$20
	Graduate	$32	$26
Black	High School	$14	$11
	College	$21	$21
	Graduate	$29	$23

Role of Conditioning

- By conditioning on available information, we can make forecasts more accurate.
- Conditioning reduces the *risk* of the forecast.
- Ignoring estimation, conditioning on more information is always better in the sense of reducing risk.