In this analysis, I will be exploring procedures for estimating demand functions. A demand function shows the causal relationship between the quantity demanded for a product and various independent variables (i.e. factors which are believed to influence demand). More specifically, I will be exploring the causal relationship between price changes for a given product and its impact on demand.

Economist theorize that holding everything else constant, when the price of a product falls, the quantity demanded of the product will increase. While, when the price of a product rises, the quantity demanded of the product will decrease. This means that there is an inverse relationship between price and quantity demanded;

$$P \downarrow \Rightarrow Q_D \uparrow$$

$$P \uparrow \Rightarrow Q_D \downarrow$$

Why might demand be downward sloping? There are a few reasons: (1) Intuitive reasoning; people would like to buy more units at lower prices and vice versa. (2) Diminishing marginal utility; consumer gets less and less additional satisfaction from consuming equal additional units of the same good. (3) Substitution effect; if the price of one good falls relative to its substitutes, then consumers will buy more of it. (4) Income effect; as the price of a good falls, the consumer can afford more units of that good, therefore their purchasing power increases.

Economists use *price elasticity* of demand to measure the responsiveness, or elasticity, of the quantity demanded of a good or service to a change in its price when nothing but the price changes. The *price elasticity* of demand can be measured by;

[1.0] $$E_P^D = \frac{\%\Delta Q_D}{\%\Delta \text{P}} = \frac{P}{Q_D} * \frac{\partial Q_D}{\partial P}$$

Since the *price elasticity* of demand is the percentage change in quantity demanded for a percentage change in price, above can equivalently be expressed as the slope of the relationship between the natural logs of Q_D and P. That is;

[1.1] $$E_P^D = \frac{ln(Q_D)}{ln(\text{P})}$$

Price elasticity of demand will always be a negative number. As price increases, quantity demanded decreases and as price decreases, quantity demanded increases.

Elastic results occur when the percentage change in quantity demanded exceeds the percentage change in price. so that the *price elasticity* is greater than 1 in absolute terms;

$$\%\Delta Q_D > \%\Delta P \Rightarrow |E_P^D| > 1$$

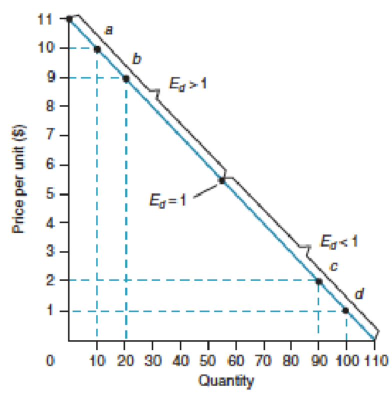
Inelastic results occur when the percentage change in quantity demanded is less than the percentage change in prices so that the *price elasticity* is less than 1 in absolute value;

$$\%\Delta Q_D < \%\Delta P \Rightarrow |E_P^D| < 1$$

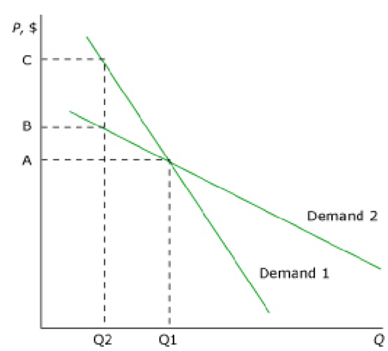
Some examples of *inelastic* goods are gas, salt, bread, tap water (necessary items), or cigarettes (addictive goods). Examples of *elastic* goods are shell gas (because consumers will shop at another gas stations), Apple (specific brands of some good), or a vacation plane ticket.

The determinants of demand elasticity are: (1) substitutability; products with close substitutes have more *elastic* demand (e.g. price of orange juice goes up, consumers substitute apple juice, cranberry juice, etc.). (2) generality; *price elasticity* for narrowly defined specific brands of products is greater than that for broadly defined product categories (e.g. demand elasticity of coke > demand elasticity of colas > demand elasticity of soft drinks). (3) proportion of income; the higher the price of the good relative to income, the greater the elasticity and vice versa. (4) degree of need; *price elasticity* of luxury goods is greater than that of necessity goods (e.g. demand elasticity for vitamins > demand elasticity for insulin). (5) complementarity; products with many complements have less *elastic* demand. (6) time; the more time that passes, the greater the demand elasticity of the product (e.g. demand elasticity for gas in the long run > demand elasticity for gas in the short run).

Something that might not be as obvious is that *price elasticity* of demand must be measured at a particular point on the demand curve;



Looking at a linear demand curve, as we move along the curve, $\frac{\Delta Q_D}{\Delta P}$ is constant, but *P* and *Q* will change. Note, that the steeper the demand curve, the more *inelastic* the demand for the good. The flatter the demand curve, the more *elastic* the demand for the good. Thus, demand 2 (from below) is relatively more *elastic* than demand 1;



One reason why *price elasticity* is important is that it tells us how much revenue changes as you change price. Suppose there is *elastic* demand ($|E_P^D| > 1$); Therefore, when price increases revenue decreases (decrease in *Q* is bigger than increase in *P*). However, when price decreases revenue increases (increase in *Q* is bigger than decrease in *P*). Suppose now that there is *inelastic* demand ($|E_P^D| < 1$); Therefore, when price increases revenue increases (decrease in *Q* is smaller than increase in *P*). However, when price decreases revenue decreases (increase in *Q* is smaller than decrease in *P*).

For the remainder of this analysis I will be covering the *logit model* for estimating demand. I will then be discussing the use of *instrumental variables* (*IV*) to deal with the endogeneity problem in my demand function– namely, when the price regressor is correlated with the error term (ε).

The reason why I have decided to use the *logit model* in estimating demand is because it is the most popular model by economic researchers. The *logit model* is the most popular model because the formula for the choice probabilities takes a closed form and is easily interpretable.

I will begin by first expressing that a consumer's indirect utility function is assumed to have the form;

[2.0]
$$u_{ijt} = \beta x_{jt} - \alpha p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

Where each consumer *i* can choose from *j* products and where markets are indexed by *t* and p_{jt} denotes the price of the product. Thus, each consumer chooses one of the following $J + 1$ mutually exclusive alternative products indexed by $j = 0, 1, \ldots, J$. The *0* choice is called the *outside option*—it means that you choose none of the above. Each *J* alternative is associated with a $K \times 1$ vector of observed characteristics ($x_{jt} = [x_{jt1} \ldots x_{jtK}]$). Each product-market *jt* is also associated with some unobserved product characteristics, ξ_{jt}. The idiosyncratic heterogeneity vector, ε_{it}, is distributed *iid* across *i* and *t*.

The ε's are distributed according to the extreme value type 2 distribution;

$$F(\varepsilon) = e^{-e^{-\varepsilon}}$$

This is a very helpful assumption, as it allows for the aggregate shares to have an analytical form. Thus, the type 2 extreme value distribution gives me the market share for product *j* in market *t*;

[2.1]
$$s_{jt}(x, \beta, \alpha, \xi) = \frac{e^{\beta x_{jt} - \alpha p_{jt} + \xi_{jt}}}{\sum_{k=0}^{J} e^{\beta x_{kt} - \alpha p_{kt} + \xi_{kt}}}$$

Where s_{jt} is the market share for each product-market *jt*. I can derive cross price derivatives as $s_j * s_k$, and own price derivative as $s_j * (1 - s_j)$.

If I normalize the mean utility associated with the outside option to 0, I get;

$$u_{i0t} = \varepsilon_{i0t}$$

Then the probability of choosing the outside good is;

[2.2]
$$s_{jt}(x, \beta, \alpha, \xi) = \frac{1}{\sum_{k=0}^{J} e^{\beta x_{kt} - \alpha p_{kt} + \xi_{kt}}}$$

And the probability of choosing any other good *j* is;

$$[2.3] \qquad s_{jt}(x, \beta, \alpha, \xi) = \frac{e^{\beta x_{jt} - \alpha p_{jt} + \xi_{jt}}}{\sum_{k=0}^{J} e^{\beta x_{kt} - \alpha p_{kt} + \xi_{kt}}}$$

By taking the log odds ratio, I can get the following linear equation;

$$[2.4] \qquad log(S_{jt}) - log(S_{0t}) = \beta x_{jt} - \alpha p_{jt} + \xi_{jt}$$

Therefore, I can estimate this via linear regression.

As mentioned earlier, ξ_{jt} are unobserved product-market level characteristics in the model. This means that there could be a list of product attributes that are unobserved in the estimation of my model (e.g. quality or prestige characteristics). For example, there can be a positive association found between price and quality that could easily be a result of unobserved self-selection. In other words, individuals who are particularly prone to quality of healthcare may be especially likely to enroll in a specific health plan (regardless of the price). ξ_{jt} can also be in the form of measurement error in prices. Since the unobserved characteristics are not independent from price, most often prices are endogenous.

The following regression example identifies where the endogeneity in *Price* can exist. Suppose;

$$[3.0] \qquad \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{\varepsilon}$$

where *X* is the question predictor variable *Price* and *Y* is demand. If I manipulate equation [3.0] by using covariance algebra (take covariance *X* throughout the equation), I get the following;

$$[3.1] \qquad \begin{aligned} cov(\hat{Y}, X) &= cov(\hat{\beta}_0 + \hat{\beta}_1 X + \hat{\varepsilon}, X) \\ . \qquad &= \hat{\beta}_1 cov(X, X) + cov(\hat{\varepsilon}, X) \\ . \qquad &= \hat{\beta}_1 var(X) + cov(\hat{\varepsilon}, X) \\ \sigma_{yx} &= \hat{\beta}_1 \sigma_X^2 + \sigma_{\varepsilon x} \end{aligned}$$

Now I will divide the population variance of the question predictor *X* throughout equation [3.1];

$$[3.2] \qquad \frac{\sigma_{yx}}{\sigma_x^2} = \hat{\beta}_1 + \frac{\sigma_{\varepsilon x}}{\sigma_x^2}$$

The $\frac{\sigma_{\varepsilon x}}{\sigma_x^2}$ ratio in equation [3.2] will give me the size of the residuals that leads to bias induced upon the $\hat{\beta}_1$ coefficient. The larger the ratio, the larger the bias (leading to the endogeneity issue). The reduction in the residual term related to bias is only possible when the second term on the right-hand side of the equation is zero;

$$\frac{\sigma_{yx}}{\sigma_x^2} = \hat{\beta}_1, \qquad when\ \sigma_{\varepsilon x} = 0$$

Statistical patterns detected in observational data do not, on their own, provide evidence of casual relationships (as illustrated in equations [3.0] – [3.2]). In this analysis, I am not relying on investigatordesigned experiment's or natural experiments to provide the required exogenous variation—instead I will be leveraging other methods to solve the endogeneity problem in *Price*. For instance, it might be possible to locate and carve out an exogenous part of the variability in the potentially endogenous predictor *Price* and use it to estimate the impact it has on demand. This can be achieved using an innovative and flexible technique called *Instrument Variables Estimation* (*IVE*).

Although the application of *IVE* to real data can be complex, the key idea is straightforward. First, note that observed differences among participants in the values of the question predictor (such as *Price*) may conceal an unknown mixture of endogenous and exogenous variation. Sometimes it is possible to carve out a part of this variation that is arguably exogenous and use only it in the estimation of causal impacts on an outcome.

In addition to these two variables, I must also have data for each participant on a special kind of background variable that is called an *instrument*. By integrating this *instrument* in a particular way onto the analysis, I can identify exogenous variation that is present in the question predictor, (e.g. *Price*) and use only it to obtain an asymptotically unbiased estimate of the causal impact of the question predictor on an outcome (e.g. demand).

The dataset I will be using to estimate a demand function will be based on the Kaiser Permanente dataset provided by the Predictive Analytics team within the Membership, Market, Sales, and Analytics department (MMSA). The variables of interest are the question predictor, *Price* (*KP Rate*) and the outcome predictor, *Subscribers*.

In the top table of figure [1], below, I have presented univariate statistics on the two key variables in this analysis. The outcome variable *Subscribers* measures the count of subscribers in each group at a point in time. My principal question predictor, *Price* (*KP Rate*), measures the average rate charged to each employer per subscriber. The average *Price* and *Subscribers* per each group in my dataset is $459 and 326, respectively.

In the middle table of figure [1], I have displayed a sample correlation matrix that summarizes the bivariate relationship between *Price* and *Subscribers*. Although the magnitude of the sample bivariate correlation between these variables is quite small, it is both statistically significant and negative (precisely the relationship direction I was anticipating), indicating that groups who have a higher *Price* tend to have a lower count of *Subscribers*.

Finally, in the bottom table of figure [1], I have presented an ordinary least-squares (OLS) regression fit, for the same outcome/predictor relationship (*Price* and *Subscribers*). I took the log of both *Price* and *Subscribers*, to evaluate the *price elasticity* relationship between the two variables. As I expected, the elasticity of *Price* on *Subscribers* is negative and statistically significant. The elasticity coefficient of *Price* is very small (abs(0.43)) indicating that *Price* is *inelastic*. However, I know that most likely *Price* is endogenous, because *Price* was not administered randomly and exogenously to groups. Since there are different prices, for each group, in unobserved ways, the relationship between *Price* and *Subscribers* in this dataset, could have easily been due to unobserved influences that are omitted in the OLS regression analysis. This means that the question predictor and residuals may be correlated, and the resulting OLS estimate regression slope may be bias. Nevertheless, having this naïve OLS estimated summary of the observed relationship is a good place to start.

Figure 1: Table results of Univariate statistics, Correlations, and Regression. Based on the endogenous variable, Price (KP Rate), and dependent variable, Subscribers

Univariate Statistics

Variable	*Mean*	*Stdv*
KP Rate	459.27	91.19
Subscribers	326	448

Sample Bivariate Correlations

	Log (KP Rate)	Log (Subscribers)
Log (KP Rate)		-0.057***
Log (Subscribers)	-0.057***	

Computed correlation used pearson-method with listwise-deletion.

OLS Regression Analysis

	Log (Subscribers)		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.60	6.17 – 9.03	**<0.001**
Log (KP Rate)	-0.43	-0.67 – -0.20	**<0.001**
Observations	4095		
R^2 / adjusted R^2	0.003 / 0.003		

I do not need to conduct a full-blown regression analysis to obtain the OLS estimate of *Log*(*Price*) on *Log*(*Subscribers*) slope in the bottom panel of figure [1]. With a single predictor in the regression model, an OLS-estimate of slope can be obtained directly by dividing the sample covariance of the outcome and question predictor by the sample covariance of the question predictor;

[4.0]

$$\begin{aligned}\hat{\beta}_1^{ols} &= \frac{\sigma_{yx}}{\sigma_x^2} \\ &= \frac{\sigma_{Log(Subscribers),Log(Price)}}{\sigma^2_{Log(Subscribers)}} \\ &= \frac{-0.0157}{0.0362} \\ &= -0.433\end{aligned}$$

Notice that the estimate in equation [4.0] is identical to that obtained in the OLS regression analysis in the bottom panel of figure [1]. However, taking this extra step of calculating the slope estimator in equation [4.0] will soon provide insight into the functioning of the OLS slope estimator itself and the *IVE*.

I know from statistical theory that if a predictor and residuals are uncorrelated, then an OLS estimate of the slope coefficient will be an unbiased estimate of the population relationship. On the other hand, if the predictors and residuals are correlated, then an OLS estimate of slope will be biased (see equations [3.0] – [3.3]). So, to accept a value of – 0.433 as an unbiased estimate of the impact of *Log*(*Price*) on *Log*(*Subscribers*) in my dataset example, I must be convinced that a groups *Price* is truly independent of the residuals in the statistical model. This would certainly be the case if a groups *Price* had been assigned randomly to groups. But, *Price*'s are determined by several factors and far from random (e.g. risk profile of group, competitive positioning, broker relationship, etc.). In summary, when a question predictor like *Price* is potentially endogenous, I cannot rely on standard OLS methods of estimation to provide an unbiased estimate of its causal impact on an outcome. Instead, I need to use a different approach.

In the top graphic (panel [a]) of figure [1], below, I have presented a Venn diagram that is useful for thinking about variation in, and covariation between, variables in either the sample or the population. I use it to marshal arguments about the variation and covariation of outcome Y and question predictor X in the population. For instance, in panel [a] of figure [2], I display a pair of intersecting ellipses. The total area of the upper ellipse represents the population variability in the outcome σ_Y^2. Similarly, the total area of the lower ellipse represents the population variability in the question predictor σ_X^2.

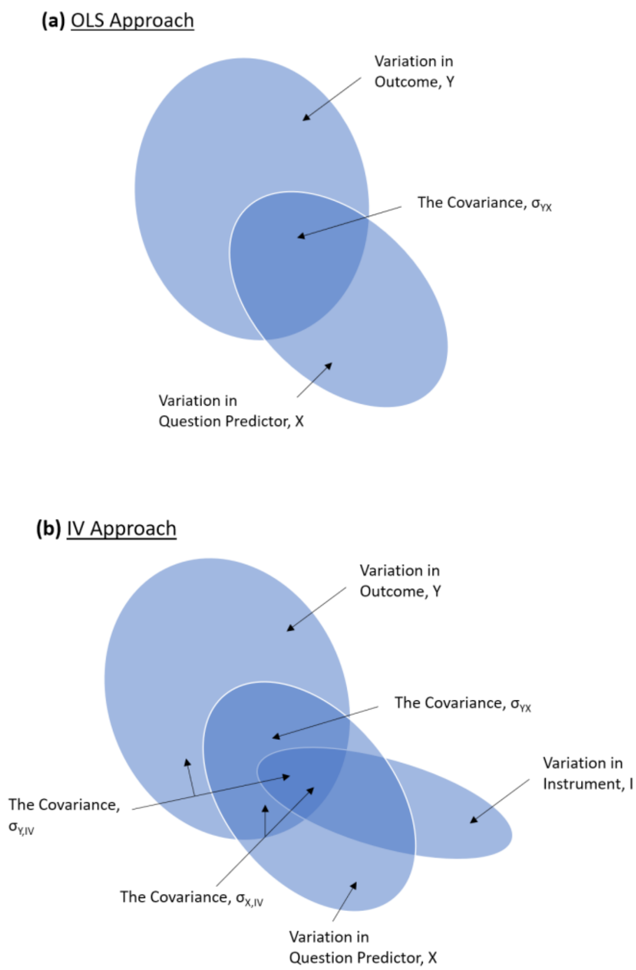
The intersection between the two ellipses symbolizes the population covariance of outcome variable and question predictor, σ_{yx}. In the analogy, when the outcome and predictor are strongly related, their covariance σ_{yx} is large. When the outcome and predictor are weakly related or not related at all, their covariance σ_{yx} will be small or zero. Conceptually, the ratio of the area of intersection to the total area of the upper ellipse then represents the portion of the total variability in outcome Y that has been successfully predicted by question predictor X (it is estimated, in the sample, by the R^2 statistic). Consequently, on the Venn diagram, the population regression slope, β_1, is represented by the area of intersection of the two ellipses (representing σ_{yx}) divided by the total area of the lower ellipse (representing σ_X^2). Working with this visual analogy for variation and covariation between outcome and predictor provides a useful explanatory tool.

There is an obvious, yet important condition embedded in the algebraic statement that underpin the OLS estimator that must hold if OLS estimation is to succeed. I mention it here because it is the precursor to an analogous condition that must be satisfied for the *IVE* technique to be effective. In equation [3.2], notice that the population variance of the question predictor, σ_X^2, which appears in the denominators of both quotients on either side of the equal sign, cannot be zero. If it were zero, and I divided throughout the equation by it, the resulting quotients would be infinite and the value of the population regression slope, β_1, inestimable. From a logical perspective, this makes sense. Why? Because it is just another way of stating that you can't detect a relationship between an outcome and a question predictor if there is no variation in the predictor. In the visual representation of figure [2], panel [a], no variation in X, would correspond to the area of the lower ellipse to shrivel up to nothing.

In addition to needing variability in X, the presence of the population variance of the question predictor σ_X^2 in the denominator of the second term to the right of the equal sign in equation [3.2] is also important. As I have implied earlier, this quotient represents the bias that will be introduced into an OLS estimate of the population slope if the question predictor and the residuals are correlated, in the population. Notice that the bias term is again a quotient: the covariance of predictor and residual, divided by the population variance of the predictor. This suggest that the magnitude of the bias that could be obtained in an OLS estimate of slope is sensitive not only to any covariation that maybe present between question predictor and residual but also to the amount of variation present in the predictor itself.

There will be a perfect storm when predictor and residuals are correlated (so that the sample estimate of numerator $\sigma_{\varepsilon x}$ is non-zero) and there is little variation in the predictor (so that the sample estimate σ_X^2 also approaches zero). In this case, the impact of any bias present due to the covariation of predictor and residuals in the numerator will be inflated, in the quotient, by the presence of the very small quantity that is present in the denominator. Clearly, this is another good reason to design research to ensure that you have substantial variability in the question predictor.

Figure 2: The use of Venn diagrams to illustrate the population variation and covariation among outcome Y, potentially endogenous question predictor, X, and instrument I, used for distinguishing the OLS and IV approaches.



In the current example, using observational data, I am estimating a demand function. My outcome variable *Subscribers*, is a measure of the number of subscribers within a group account enrolled in a Kaiser Healthcare plan. The important question predictor, *Price*, is used to calculate the *price elasticity* for each group account. Theoretically, there should be a causal relationship between the latter and the former. Consequently, I would like to use the observational data to obtain a credible estimate of the causal impact of *Price* on *Subscribers*. However, I suspect that the question predictor *Price* is potentially endogenous because Kaiser Permanente has been able to choose predetermined rates for each group account. As a result, an OLS estimate of the relationship between *Subscribers* on *Price* may provide a biased view of the hypothesized underlying population relationship between the marked price and the count of subscribers within a group.

In statistics, as in life, it is usually the case that we can always do better if we have some way to incorporate additional useful information into our decision. Setting all skepticism aside, let's imagine for a moment that I had information available on an additional and very special kind of variable—named *IV*, for *instrument variable*—that has also been measured for all groups in the sample dataset. Now ask ourselves: What properties would such an *IV* need to have, to be helpful? How could I incorporate it into the analysis, if I wanted to end up with an unbiased estimate of the critical relationship between *Subscribers* and *Price*?

Although this seems to be a completely inhospitable analytic situation, I can gain insight by again applying covariance algebra to the hypothesized population regression featured in equation [3.0]. This time, though, instead of taking the covariance throughout with predictor *X* (as I did in equation [3.1]), I will take them with the new *IV*. This leads to the following result;

[5.0]
$$\begin{aligned} cov(Y, IV) &= cov(\hat{\beta}_0 + \hat{\beta}_1 X + \hat{\varepsilon}, IV) \\ &= \hat{\beta}_1 cov(X, IV) + cov(\hat{\varepsilon}, IV) \end{aligned}$$

Or, more parsimoniously;

[5.1]
$$\sigma_{Y,IV} = \hat{\beta}_1 \sigma_{X,IV} + \sigma_{\varepsilon,IV}$$

Dividing through by the population covariance of *X* and *IV*;

[5.2]
$$\frac{\sigma_{Y,IV}}{\sigma_{X,IV}} = \hat{\beta}_1 + \frac{\sigma_{\varepsilon,IV}}{\sigma_{X,IV}}$$

Here, surprisingly, notice a second interesting consequence of the specification of the population linear regression model. From equation [5.2], the population covariance of *Y* with *IV* ($\sigma_{Y,IV}$) divided by the population covariance of *X* with *IV* ($\sigma_{X,IV}$) is again equal to the critical parameter representing the key population relationship of interest ($\hat{\beta}_1$), provided that the second term on the right-hand side of the equation is zero. And it is zero when the new instrument is uncorrelated with the residual in the population regression model. That is;

[5.3]
$$\frac{\sigma_{Y,IV}}{\sigma_{X,IV}} = \hat{\beta}_1, \qquad when\ \sigma_{\varepsilon,IV} = 0$$

In other words, if there were some way to locate an *IV* that is incontrovertibly uncorrelated with the population residuals in the question model, I would be home free. Then, the slope of the causal relationship between *Price* on *Subscribers* in the population would simply be equal to the ratio of two important population covariances: (a) the covariance of the outcome and instrument, $\sigma_{Y,IV}$, and (b) the covariance of question predictor and instrument, $\sigma_{X,IV}$. So, if *IV* were known and its values are measured in the sample, I could simply estimate each of these respective covariances by their corresponding sample statistics and replace them in the quotient by their sample equivalents. This would provide me with an (asymptotically) unbiased estimate of the causal impact of *Price* on *Subscribers* in the population. This alternative estimator is called the *instrument variable estimator* of β_1;

[5.4]
$$\hat{\beta}_1^{IVE} = \frac{\sigma_{Y,IV}}{\sigma_{X,IV}}$$

Because statisticians refer to covariance as one of the *second moments* of a bivariate distribution, the expression on the right-hand side of equation [5.4] is referred to the *method-of-moments IVE* of β_1.

In figure [3], below, I have reproduced the previous naïve findings (from figure [2]) and also have introduced the variable that I claim works well as an *IV* in the investigation of the causal impact of *Price* on *Subscribers*. This variable is *Risk Scores*. *Risk Scores* is a continuous variable that calculates the risk profile of members (e.g. healthy vs sick, high utilizers vs low utilizers, costly members vs inexpensive members, etc.). Notice that, on average, groups had an average *Risk Score* of 1.59, but the sample standard deviation is almost as large as the mean, indicating that there is considerable variability in the values of the potential instrument.

I have to argue that the instrument variable *Risk Scores* is unlikely to be correlated with the residuals in the regression of outcome *Subscribers* on the potentially endogenous variable *Price* (more on this later). This assumption might be difficult to hold. *Risk Scores* is based off the health profile of members. It could be argued that locations with higher unemployment rates and less access to healthier food can impact *Risk Scores*. Later, I include control variables to avoid this potential violation. However, for illustration purposes, suppose for now that *Risk Scores* is not correlated with the residuals in the regression of outcome *Subscribers* on the potentially endogenous variable *Price*. Thus, although some part of the variation in *Price* may have been determined endogenously, some other parts of it may be exogenous and related to *Risk Scores*. If this is truly the case, then I can employ *Risk Scores* as an instrument to obtain an asymptotically unbiased estimate of the causal relationship between *Price* and *Subscribers*.

In middle table of figure [3] (sample bivariate correlations), notice that the endogenous question predictor *Price* and instrument *Risk Scores* are indeed related. This relationship is very important. The higher the *Risk Score*, the higher the *Price* (*KP Rate*)—which make intuitive sense. In addition, the outcome *Subscribers* has a negative relationship with the instrument *Risk Score*. Thus, the greater the *Risk Score* the less *Subscribers* enroll within a group account. Substituting the corresponding sample covariances into equation [5.4], I obtain an asymptotically unbiased *method-of-moments IVE* of the impact of *Price* on *Subscribers*;

$$\begin{aligned}\hat{\beta}_1^{IVE} &= \frac{\sigma_{Y,IV}}{\sigma_{X,IV}} \\ &= \frac{\sigma_{Log(Subscribers),Risk\ Score}}{\sigma_{Log(KP\ Rate),Risk\ Score}} \\ &= \frac{-0.2716}{0.0516} \\ &= -5.2595\end{aligned}$$

Notice that this coefficient is still negative but much larger in magnitude (- 5.2595 vs - 0.433). This suggest that groups are much more price sensitive to price changes then initially estimated. Provided that the instrument, *Risk Score*, satisfies the critical assumption I described earlier, then this new value of - 5.2595 is an asymptotically unbiased estimate of the impact of *Price* on *Subscribers*.

Figure 3: Table results of Univariate statistics, Correlations, and Methods-of-Moments IVE Estimate. Based on the endogenous variable Price (KP Rate), IV Risk Score, and dependent variable, Subscribers

Univariate Statistics

Variable	*Mean*	*Stdv*
KP Rate	459.27	91.19
Subscribers	326	448
Risk Score	1.59	1.39

Sample Bivariate Correlations

	Log (KP Rate)	Log (Subscribers)	Risk Score
Log (KP Rate)		-0.057***	0.196***
Log (Subscribers)	-0.057***		-0.135***
Risk Score	0.196***	-0.135***	

Computed correlation used pearson-method with listwise-deletion.

Method-of-Moments IVE Estimate

	1st Stage Log (KP Rate)			**2nd Stage Log (Subscribers)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	6.07	6.06 – 6.08	**<0.001**	37.09	29.84 – 44.34	**<0.001**
Risk Score	0.03	0.02 – 0.03	**<0.001**			
Log $(\widehat{\text{KP}}$ Rate)				-5.26	-6.45 – -4.07	**<0.001**
Observations	4095			4095		
R^2 / adjusted R^2	0.038 / 0.038			0.018 / 0.018		

It is useful to explore the logic upon which this new method of estimation is based. Conceptually, during the *IVE* process, I used my instrument—which I regard as exogenous by assumption (and therefore uncorrelated with the residuals in the main regression model)—to carve out part of the variation in the question predictor that is also exogenous and then I use only the latter part in the estimation of the regression slope. I can illustrate this statement by extending the earlier graphical analogy to the lower panel (panel [b]) of figure [2]. In the new panel, I have replicated the original Venn diagram in the upper panel (panel [a]), with the same pair of overlapping ellipses representing the variances and covariances of the outcome and question predictor, as before. Then, across the two intersecting ellipses, I have overlaid a third ellipse (much narrower than the first two) to represent variation in the *IV*. Notice that the latter ellipse has been drawn to overlap both the first two ellipses, thereby co-varying uniquely with both. However, it does not overlap any of the original residual variation in *Y* (which is the part of the upper ellipse that falls beyond the reach of variation in the question predictor *X*.

I have drawn the new figure like this because, by definition, a successful *IV* must not be correlated with those residuals, and therefore can be no overlap of instrument and residual variation. Finally, in panel [b] of figure [2], I suggest that there may be some substantial part of the *IV*'s variation that is independent of variation in the question predictor; this is why I have drawn the *IV* ellipse to stick out to the right of the question predictor, *X*.

When I carry out successful *IVE*, it is as though I have allowed the *IV* ellipse (in figure [2], panel [b]) to carve out the corresponding parts of the original *X* (the question predictor) ellipse and the intersection of *Y* on *X*. And, because variation in the *IV* is exogenous (by an assumption that I still need to defend), the parts that I have carved out must also be exogenous. Then in forming the *IVE*, I restrict implicitly to work with only the variation in outcome and question predictor that is shared (i.e. the intersects or covariates) with the new *IV* within the *IV* ellipse. Within this shared region, I again form a quotient that is a ratio of a "part" to a "whole" to provide the new *IVE* of the *Y* on *X* slope. This quotient is the ratio of the covariation shared between outcome and instrument to the covariation shared by question predictor and instrument, respectively. I can identify the corresponding region in figure [2], panel [b]. They are the regions where the three ellipses intersect (variation in *Y*, *X*, and *IV*), and where the two ellipses—variation in *X* and *IV* intersect—respectively. In a sense I have used the instrument to carve up the old dubious variation and covariation in both outcome and question predictor into identifiable parts and have picked out and incorporated into the new estimate only those parts that I know are—by assumption—exogenous. All in all, the new *IV* estimator enacts exactly the same principles as the OLS estimator, but within the region defined by the exogenous variation of the instrument.

Notice that, when I use *IVE*, I restrict the analytic focus in one other important way. By working only within the *IV* ellipse, I have restricted myself to a region of variation in both outcome and question predictor that is smaller than the original region of variation with which the initial OLS regression analysis was conducted. Artificially limiting the variation in outcome and predictor that is being incorporated into any analysis means that the obtained estimate will usually have lower precision—that is, its standard error will be larger than the corresponding (but biased) OLS estimate. Thus, in using *IV* methods to provide an asymptotically unbiased estimate of the causal relationship of interest, I have traded away some of the original precision and statistical power, making it harder to reject the corresponding null hypothesis, for the benefit of knowing that I now have an asymptotically unbiased estimate. Unfortunately, if I choose an instrument that has very limited variation, I exacerbate this problem enormously. An instrument of limited variation can only carve out a small part of the variation in outcome and question predictor for incorporation into the *IV* estimate.

It is wise to spend some time now on the two critical assumptions *IVE* must hold. *IVE* is an approach that offers important advantages to empirical researchers who seek to draw unbiased causal conclusions from quasi-experimental or observational data. However, it is critical to keep in mind that any application *IVE* depends strongly on additional assumptions not required of OLS methods. In practice, the veracity of one of these assumptions proves relatively easy to confirm. Unfortunately, this is not the case for the second assumption.

The “easy-to-prove” condition for successful *IVE* is that the instrument must be related to the potentially endogenous question predictor. In other words, the population covariance of question predictor and instrument, $\sigma_{X,IV}$, cannot be zero;

[6.0] $$Cov(IV,X) \neq 0$$

This condition seems obvious, from both a logical and a statistical perspective. If the question predictor and instrument are unrelated, then I cannot use the instrument successfully to carve out any part of the variation in *X*, let alone any exogenous variation, and so the *IVE* will inevitably fail. Fortunately, in the case of the *price elasticity* example from above, both the *IV* (*Risk Score*) and *Price* (*KP Rate*) are related and statistically significant (see middle table of figure [4]).

The second important condition that must be satisfied for successful *IVE* is that the instrument cannot be related to the unobserved effects (the residuals). In other words the covariance of *IV* and residuals, $\sigma_{\varepsilon,IV}$, must be zero;

[6.1] $$Cov(IV,\varepsilon) = 0$$

This was shown throughout the algebraic development that led to the *IV* estimator in equations [5.2] and [5.3]. The condition is also appealing logically. If the instrument were correlated with the residuals, it would suffer from the same problem as the question predictor itself. Thus, it could hardly provide a solution to the endogeneity problem. Although this condition is more difficult to provide, I will argue that it could hold based on additional control variables correlated with the *IV* (*Risk Scores*) I used from the dataset.

Earlier, in this paper, I mentioned that *Risk Scores* is based off the health profile of members. It could be argued that locations with higher unemployment rates and less access to healthier food can impact *Risk Scores*. Therefore, to account for potential violations of equation [6.1], I have included a series of control variables to help hold my argument that the instrument cannot be related to the residuals.

The fist control variable I included is *Family Content Ratio* (*FCR*). Notice that in the middle table of figure [5] (sample bivariate correlations), *FCR* is both positively correlated and statistically significant with *Price* (*KP Rate*). *FCR* is a continuous variable that measures the average count of dependent coverage for each enrolled subscriber. The top table in figure [5] (univariate statistics), indicates on average, subscribers have at least one other dependent covered by their Kaiser healthcare plan (e.g. spouse, child, etc.). *FCR* is heavily negatively correlated with the *IV* variable, *Risk Scores*, indicating the higher the *FCR* the lower the *Risk Score*.

The second control variable I included is Proportion of Black or Hispanic members (*Prop. Black & Hisp*) enrolled in a Kaiser health care plan. *Prop. Black & Hisp* is negatively correlated with *Price*. This indicates the higher the proportion of Black and Hispanic members enrolled the lower the *Price*.

Finally, the last control variable I included is *Household Income*. *Household Income* is a continuous variable that measure the average *Household Income* of all employees within a group account. This last variable is positively correlated with *Price*. *Household Income* also has the largest correlation with *Price* over all other control variables. I believe *Household Income* is an important control variable because people with higher incomes tend to have access to better food.

Figure 4: Table results of Univariate statistics, Correlations, and Methods-of-Moments IVE Estimate. Based on the endogenous variable Price (KP Rate), IV Risk Score, FCR, Prop. Black & Hisp, Household Income, and dependent variable Subscribers

Univariate Statistics

Variable	*Mean*	*Stdv*
KP Rate	459.27	91.19
Subscribers	326	448
Risk Score	1.59	1.39
FCR	2.02	0.46
Prop. Black & Hisp	0.35	0.15
Household Income	90,947.02	16,013.09

Sample Bivariate Correlations

	Log (KP Rate)	Log (Subscribers)	Risk Score	Household Income	Prop. Black & Hisp	FCR
Log (KP Rate)		-0.057***	0.196***	0.113***	-0.100***	0.082***
Log (Subscribers)	-0.057***		-0.135***	-0.031*	0.308***	0.001
Risk Score	0.196***	-0.135***		-0.026	-0.088***	-0.315***
Household Income	0.113***	-0.031*	-0.026		-0.278***	0.006
Prop. Black & Hisp	-0.100***	0.308***	-0.088***	-0.278***		-0.003
FCR	0.082***	0.001	-0.315***	0.006	-0.003	

Computed correlation used pearson-method with listwise-deletion.

Method-of-Moments IVE Estimate

	1st Stage Log (KP Rate)			**2nd Stage Log (Subscribers)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	5.83	5.78 – 5.88	**<0.001**	25.29	19.55 – 31.04	**<0.001**
Risk Score	0.03	0.03 – 0.04	**<0.001**			
Household Income	0.00	0.00 – 0.00	**<0.001**	0.00	0.00 – 0.00	**<0.001**
Prop. Black & Hisp	-0.06	-0.10 – -0.02	**0.002**	2.73	2.43 – 3.03	**<0.001**
FCR	0.07	0.05 – 0.08	**<0.001**	0.13	0.03 – 0.23	**0.010**
Log ($\widehat{\text{KP}}$ Rate)				-3.67	-4.63 – -2.71	**<0.001**
Observations	4095			4095		
R^2 / adjusted R^2	0.078 / 0.077			0.110 / 0.109		

Including the control variables in my estimation of *price elasticity* of demand (*Log ($\widehat{KP\ Rate}$)*) dropped the original elasticity value from abs(5.26) to my current elasticity value of abs(3.67). This reinforces the importance of the discussion around my assumption in equation [6.1]. Endogeneity may still exist in *Price*, therefore caution and skepticism is heavily recommended.

Using the chosen *IV* variable (*Risk Score*) with a series of control variables (*FCR, Prop. Black & Hisp, Household Income*), I will estimate the demand function in equation [2.4] to calculate the *price elasticity* of demand for each group in my dataset. Prior to calculating the demand function I need to define S_{jt} and S_{0t} from the dataset;

$$S_{jt} = Subscribers / Eligible$$

and,

$$S_{0t} = (Eligible - Subscribers) / Eligible$$

Therefore, estimating equation [2.4] equals,

[7.0] $$log(S_{jt}) - log(S_{0t}) = \alpha\widehat{Price} + \beta_1 Household\_Income + \beta_2 Prop\_BlackHisp + \beta_3 FCR + \varepsilon$$

[7.1] $$\alpha\widehat{Price} = \beta_1 Rate\_Increase + \beta_2 Household\_Income + \beta_3 Prop\_BlackHisp + \beta_4 FCR + \varepsilon$$

The model output of equation [7.0], is in the figure below;

Method-of-Moments IVE Estimate

	1st Stage Log (KP Rate)			2nd Stage Log (S_{jt}) – Log (S_{0t})		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	328.666	304.79 – 352.55	**<0.001**	2.704	2.00 – 3.41	**<0.001**
Risk Score	16.475	14.42 – 18.53	**<0.001**			
Household Income	0.000	0.00 – 0.00	**<0.001**	-0.000	-0.00 – -0.00	**<0.001**
Prop. Black & Hisp	-26.520	-44.78 – -8.26	**0.004**	-0.423	-0.67 – -0.18	**0.001**
FCR	30.421	24.25 – 36.60	**<0.001**	0.256	0.18 – 0.34	**<0.001**
Log ($\widehat{KP}$ Rate)				-0.007	-0.01 – -0.01	**<0.001**
Observations	4095			4095		
R^2 / adjusted R^2	0.078 / 0.077			0.062 / 0.061		

To calculate *price elasticity* of demand for each group use the following formula;

[7.2] $$Price\ Elasticy = \alpha * Price_{jt} * (1 - S_{jt})$$