4

PERSISTENCE ANALYSIS

Many of the variables in empirical asset pricing research are intended to capture persistent characteristics of the entities in the sample. This means that the characteristic of the entity that is captured by the given variable is assumed to remain reasonably stable over time. Such variables are frequently estimated using historical data, and the value calculated from the historical data is assumed to be a good estimate of the given characteristic for the entity going forward. For example, the value of a stock's beta from the Capital Asset Pricing Model (Sharpe (1964), Lintner (1965), Mossin (1966)) is generally assumed to be a persistent characteristic of the stock, and it is frequently estimated from regressions of the stock's returns on the returns of the market portfolio using historical data. This is exactly how our variable β is calculated.

In this chapter, we discuss a technique that we call persistence analysis. We use persistence analysis to examine whether a given characteristic of the entities in our sample is in fact persistent. Persistence analysis can also be used to examine the ability of the variable in question to capture the desired characteristic of the entity. The basic approach is to examine the cross-sectional correlation between the given variable measured at two different points in time. If this correlation is high, this indicates that the variable is persistent, whereas low correlations indicate little or no persistence. This technique is not as widely used in the empirical asset pricing literature as the other techniques presented in Part I. We discuss it here and use it throughout this text because one of the objectives of this book is to provide a thorough understanding of the variables most commonly used throughout the empirical asset pricing literature.

Empirical Asset Pricing: The Cross Section of Stock Returns, First Edition.
Turan G. Bali, Robert F. Engle, and Scott Murray.

4.1 IMPLEMENTATION

As with the other methodologies presented in this text, implementation of persistence analysis is done in two steps. The first step involves calculating cross-sectional correlations between the given variable X measured a certain number of periods apart. The second step involves calculating the time-series average of each of these cross-sectional correlations.

4.1.1 Periodic Cross-Sectional Persistence

The first step in persistence analysis is to calculate the cross-sectional correlation between the variable under consideration, X, measured τ periods apart. This will be done for each time period t where both the time period t and the time period $t + \tau$ fall during the sample period. The entities used to calculate the cross-sectional correlation will be all entities i for which a valid value of the variable X is available for both period t and period $t + \tau$. For each time period t, we therefore define $\rho_{t,t+\tau}(X)$ as the cross-sectional Pearson product–moment correlation between X measured at time t and X measured at time $t + \tau$. Specifically, we have

$$\rho_{t,t+\tau}(X) = \frac{\sum_{i=1}^{n_t}[(X_{i,t} - \overline{X}_t)(X_{i,t+\tau} - \overline{X}_{t+\tau})]}{\sqrt{\sum_{i=1}^{n_t}(X_{i,t} - \overline{X}_t)^2}\sqrt{\sum_{i=1}^{n_t}(X_{i,t+\tau} - \overline{X}_{t+\tau})^2}} \quad (4.1)$$

where $\overline{X}_t$ is the mean value of $X_{i,t}$ and the summations and means are taken over all entities i for which a valid value of X is available in both periods t and $t + \tau$. n_t represents the number of such entities. Frequently, before the correlations are calculated, the values of X from month t are winsorized to remove the effect of outliers. The values of X from month $t + \tau$ are separately winsorized at the same level.

We illustrate this using β and values of τ between 1 and 5 inclusive. Our analysis will therefore examine the persistence of β measured one, two, three, four, and five years apart. Prior to calculating the cross-sectional correlations for each period t, the data are winsorized at the 0.5% level. To be perfectly clear, for each month t, we first find all entities that have valid values of β in both periods t and $t + \tau$. We then winsorize the corresponding values of β in each of the months t and $t + \tau$ separately. The annual values for these cross-sectional correlations are presented in Table 4.1. The year t is presented in the first column and the subsequent columns present the values of $\rho_{t,t+\tau}(\beta)$ for $\tau \in \{1, 2, 3, 4, 5\}$.

The results in Table 4.1 indicate that values of β measured one year apart ($\rho_{t,t+1}(\beta)$) exhibit cross-sectional correlations between 0.39 (t = 1992) and 0.80 (t = 2008). As might be expected, the correlations between β measured at longer lags τ tend to be lower than the correlations measured at shorter lags, although this is not always the case. When measured five years apart ($\rho_{t,t+5}(\beta)$), the table indicates that the correlation between β and its lagged counterpart is between 0.25 (t = 2000) and 0.56 (t = 2006). We withhold further interpretation of the results until later in the chapter.

TABLE 4.1 Annual Persistence of β

This table presents the cross-sectional Pearson product–moment correlations between β measured in year t and β measured in year $t+\tau$ for $\tau \in \{1,2,3,4,5\}$. The first column presents the year t. The subsequent columns present the cross-sectional correlations between β measured at time t and β measured at time $t+1$, $t+2$, $t+3$, $t+4$, and $t+5$.

t	$\rho_{t,t+1}(\beta)$	$\rho_{t,t+2}(\beta)$	$\rho_{t,t+3}(\beta)$	$\rho_{t,t+4}(\beta)$	$\rho_{t,t+5}(\beta)$
1988	0.50	0.48	0.47	0.39	0.34
1989	0.52	0.45	0.38	0.35	0.36
1990	0.55	0.45	0.42	0.40	0.37
1991	0.46	0.43	0.41	0.37	0.40
1992	0.39	0.37	0.36	0.41	0.38
1993	0.40	0.33	0.38	0.39	0.39
1994	0.38	0.39	0.38	0.37	0.33
1995	0.46	0.44	0.38	0.40	0.48
1996	0.53	0.48	0.43	0.52	0.53
1997	0.55	0.46	0.48	0.51	0.53
1998	0.51	0.50	0.53	0.53	0.50
1999	0.57	0.59	0.53	0.52	0.40
2000	0.79	0.58	0.56	0.58	0.25
2001	0.70	0.66	0.62	0.35	0.41
2002	0.79	0.64	0.50	0.49	0.38
2003	0.70	0.54	0.51	0.42	0.39
2004	0.62	0.60	0.45	0.38	0.34
2005	0.73	0.60	0.55	0.48	0.53
2006	0.67	0.56	0.50	0.56	0.56
2007	0.69	0.60	0.60	0.59	0.51
2008	0.80	0.69	0.64	0.60	
2009	0.73	0.65	0.59		
2010	0.76	0.70			
2011	0.79				

However, it is worth noting that for years t toward the end of the sample, in some cases the persistence values are missing. The reason for this is that, for example, in year 2009, to calculate the correlation between β measured in 2009 and β measured four years in the future ($\tau = 4$), we would need data from year 2013. As these data are not available in the version of the Center for Research in Security Prices (CRSP) database used to construct the methodology sample, we are unable to calculate this value. The reasons for the other missing entries are analogous.

4.1.2 Average Cross-Sectional Persistence

Although periodic cross-sectional persistence values such as those presented in Table 4.1 are quite informative, they are quite difficult to read and draw conclusions from. We therefore want to summarize these periodic values more succinctly. As with the other analyses we discuss, the main objective is to understand the persistence of the variable X in the average cross section. We therefore summarize the results by simply taking the time-series average of the periodic cross-sectional correlations. We denote these average persistence values using $\rho_\tau(X)$ where the subscript indicates the number of lags. Specifically, we have

$$\rho_\tau(X) = \frac{\sum_{t=1}^{N-\tau} \rho_{t,t+\tau}(X)}{N - \tau} \tag{4.2}$$

where N is the number of periods in the sample. Throughout the remainder of this book, we will refer to these values as the persistence of X at lag τ.

In Table 4.2, we present the persistence of β at lags of one, two, three, four, and five years. The results indicate that, consistent with what was observed in the annual persistence values presented in Table 4.1, the persistence of values of β measured one year apart, 0.62, is quite strong. The level of persistence drops off substantially as the amount of time between the measurement periods increases. When measured at a lag of 5 years, the average persistence of β has decreased to 0.42.

4.2 INTERPRETING PERSISTENCE

Interpreting the results of the persistence analyses is fairly straightforward. In general, a higher degree of time-lagged cross-sectional correlation in the given variable is indicative of higher persistence, although there are several caveats to this that must be understood to properly make use of this technique.

We begin our discussion of the interpretation of persistence analysis results by discussing potential causes of low persistence. Exactly what qualifies as low persistence is not perfectly well defined and depends on how long the lag is between the times of measurement (τ), how persistent the actual characteristic being captured by the variable is thought to be, and how accurately the variable is expected to capture the actual characteristic. There are generally two reasons that a variable may exhibit low or zero persistence. The first is that the characteristic being measured is in fact

TABLE 4.2 Average Persistence of β

This table presents the time-series averages of the cross-sectional Pearson product–moment correlations between β measured in year t and β measured in year $t + \tau$ for $\tau \in \{1, 2, 3, 4, 5\}$.

$\rho_1(\beta)$	$\rho_2(\beta)$	$\rho_3(\beta)$	$\rho_4(\beta)$	$\rho_5(\beta)$
0.61	0.53	0.48	0.46	0.42

not persistent. The second is that the variable used to proxy for the given characteristic does a poor job at measuring the characteristic under examination. In this case, even if the given characteristic of the entities in the sample is highly cross-sectionally persistent, the failure of the variable X to capture cross-sectional variation in this characteristic will cause the persistence analysis to generate a low value of $\rho_\tau(X)$. In this sense, the persistence analysis suffers from a dual hypothesis problem, as failure to find persistence does not necessarily indicate a lack of persistence in the characteristic under investigation. Low values of $\rho_\tau(X)$ may also indicate a failure of X to capture that characteristic. Thus, we must be careful when concluding that a certain characteristic of the entities in the sample is not cross-sectionally persistent based on the results of the persistence analysis. To reach such a conclusion, we must be highly confident that the variable X does in fact capture the characteristic under examination. On the other hand, if one is extremely confident, for reasons beyond the scope of the persistence analysis, that the characteristic in question is in fact highly cross-sectionally persistent, low values of $\rho_{t,t+\tau}(X)$ likely indicate that the variable X does a poor job at capturing cross-sectional variation in the characteristic. In the end, however, regardless of the reason for the lack of persistence in X, if X is intended to capture a persistent characteristic of a firm, but X does not exhibit persistence, then X is not a good measure of the characteristic of interest.

When the persistence analysis produces high values of $\rho_\tau(X)$, this very likely indicates both that the characteristic in question is in fact persistent and that the variable X does a good job at measuring the characteristic. There are two caveats with this statement that must be addressed. The first is that it is possible that the variable X is unintentionally capturing some persistent characteristic of the entities in the sample that is different from the characteristic that X is designed to capture. Thus, perhaps a more correct statement is that high values of $\rho_\tau(X)$ indicate that whatever characteristic is being captured by X is in fact persistent. If X does in fact capture the intended characteristic, then we can conclude that the characteristic is in fact persistent. Therefore, assuming sufficient effort has been devoted to designing the calculation of X such that it can reasonably be expected to capture the intended characteristic, high values of $\rho_\tau(X)$ are interpreted as indicating that the given characteristic is in fact persistent.

The second, and much more important, caveat associated with concluding that a characteristic is persistent is that in many cases there is a mechanical reason related to the calculation of X that would result in strong cross-sectional correlation between X_t and $X_{t+\tau}$ even if the characteristic in question is not persistent. In most cases, the reason for such a mechanical effect is that some subset of the data used to calculate X at times t and $t + \tau$ are the same. This is frequently the case when a variable is calculated from historical data covering more than τ periods. For example, if X_t is calculated using k periods of historical data up to and including period t, where $k > \tau$, then X_t and $X_{t+\tau}$ are calculated using some of the same data and are therefore likely to be correlated in the cross section as a result. For this reason, when X is calculated using k periods of historical data, persistence analysis is only effective when $\tau \geq k$.

In addition to examining whether a given characteristic of the entities in the sample is cross-sectionally persistent, persistence analysis can also be helpful in determining the optimal measurement period that should be used to calculate a given variable. Many of the variables used throughout the empirical asset pricing literature are calculated based on historical data. When calculating these variables, researchers are faced with the decision of how long a calculation period to use. Increasing the length of the calculation period means that more data are used in the calculation of the variable, which can increase the accuracy of the measurement. However, using longer calculation periods also means that when calculating the variable for time period t, data from many periods prior to t are used, and this data may no longer be reflective of the given characteristic of the entity at time t. For this reason, extending the calculation period too long may result in decreasing accuracy of measurement. How to optimally make this trade-off depends on the persistence of the characteristic being measured. While certainly none of the variables studied in asset pricing research are perfectly persistent, different variables exhibit different degrees of persistence.

To help determine the optimal calculation period for a variable calculated from historical data, we can examine the patterns in the persistence of the variable measured using different calculation periods. The main concept behind this application of persistence analysis is that the cross-sectional persistence of even the most persistent characteristic is likely to decay over time. Therefore, let us assume we have two variables X_1 and X_2 that measure the same characteristic using the same formula but applied to different calculation periods of length τ_1 and τ_2, respectively, and without loss of generality, let $\tau_1 > \tau_2$. Let us also make the assumption that X^1 and X^2 are equally accurate measures of the given characteristic.

If X^1 and X^2 are equally accurate measures of the characteristic, then based on the assumed decay in persistence as the value of τ increases, we would expect the persistence of X^1 measured at a lag of $\tau = \tau_1$ to be greater than the persistence of X^2 measured at lag $\tau = \tau_2$. Notice here that the lag at which the comparison of the persistence is done is such that neither analysis has the overlapping data issue discussed earlier. If the persistence of X_2 at lag $\tau = \tau_2$ is actually greater than that of X_1 at lag $\tau = \tau_1$, this is a contradiction of what would be expected if X_1 and X_2 were equally accurate measures. This therefore indicates that using τ_2 periods of data to calculate X provides a more accurate measure of the underlying characteristic than using τ_1 periods, as the additional amount of data used in the calculation apparently overcomes the decay in the persistence at longer lags τ.

If the persistence of X^2 at lag $\tau = \tau_2$ is less than that of X^1 at lag $\tau = \tau_1$, the results are a bit more challenging to interpret, but it can generally be taken to mean that the decay in the persistence over a period of $\tau_2 - \tau_1$ periods is substantial enough to overcome any additional benefit of using τ_2 periods of data, compared to τ_1, to calculate X. If this is the case, it may also be an indication that using a full τ_2 periods of data is too long a measurement period because the given characteristic of the firm does in fact change substantially over periods of length τ_2.

There is a practical consideration that may have an effect on using persistence to determine the optimal measurement period for the given variable X. This consideration is that the sample changes over time. The calculation of the value of $\rho_{t,t+\tau}(X)$ is done using only those entities that are in the sample at both times t and $t + \tau$. In most cases, the set of entities in the sample at both time t and time $t + \tau_1$ is likely to be a superset of the set of entities in the sample at both time t and time $t + \tau_2$ ($\tau_2 > \tau_1$). Furthermore, in many cases, the set of entities that remain in the sample until time $t + \tau_2$ is likely to be more "well-behaved" than those that do not remain in the sample, where well-behaved can be taken to mean that the calculation of the variable X is a more accurate measure of the characteristic being examined for such entities than for entities that are not well-behaved. If these not well-behaved entities are more likely to enter and then drop out of the sample over a small number of periods, it is possible that using persistence analysis to examine the quality of a variable as described in this section may be misleading. That being said, for the analyses performed in this text this is unlikely to be a substantial issue, as the number of entities (stocks in this case) in each cross section is quite large relative to the number of stocks that drop out of the sample each period.

4.3 PRESENTING PERSISTENCE

Throughout this book, we will present the results of persistence analyses by displaying the values of $\rho_\tau(X)$. Each column in the tables that present the persistence analysis results will correspond to a given variable, indicated in the first row of the column. Each row will correspond to a given value of τ.

The results of persistence analyses for each of β, *Size*, and *BM* using lags of one, two, three, four, and five years are presented in Table 4.3. The results indicate that all three variables are highly persistent. The persistence of β measured at lag of one year ($\tau = 1$) is 0.61 and that of *Size* is 0.96, and for *BM* the persistence at lag of one year is 0.74. The results for each of these variables indicate fairly strong persistence at lags of up to five years. *Size* is very highly persistent, as the average cross-sectional

TABLE 4.3 Persistence of β, *Size*, and *BM*

This table presents the results of persistence analyses of β, *Size*, and *BM*. For each year t, the cross-sectional correlation between the given variable measured at time t and the same variable measured at time $t + \tau$ is calculated. The table presents the time-series averages of the annual cross-sectional correlations. The column labeled τ indicates the lag at which the persistence is measured.

τ	β	*Size*	*BM*
1	0.61	0.96	0.74
2	0.53	0.92	0.59
3	0.48	0.90	0.50
4	0.46	0.89	0.46
5	0.42	0.88	0.43

correlation between *Size* measured five years apart is 0.88, only slightly lower than when the persistence is measured at a lag of one year. The decay in the persistence of β and *BM* is substantially more pronounced, but even after five years, β and *BM* continue to exhibit substantial persistence.

4.4 SUMMARY

In this chapter, we have presented a methodology for examining the persistence of a given variable. The methodology has two primary applications. If we assume that the variable accurately measures the characteristic that it is intended to capture, then persistence analysis can be used to examine how persistent the given characteristic is in the cross section of the entities in the sample. If we assume the characteristic that the variable is intended to measure is in fact persistent, then we can use persistence analysis to examine the accuracy with which the variable captures the given characteristic and the optimal measurement period to use when calculating the variable. Of course, no characteristic is perfectly persistent and no variable perfectly captures the characteristic it is designed to measure. Despite these caveats, persistence analysis is a useful tool that we will employ throughout this text.

REFERENCES

Lintner, J. 1965. Security prices, risk, and maximal gains from diversification. Journal of Finance, 20(4), 687–615.

Mossin, J. 1966. Equilibrium in a capital asset market. Econometrica, 34(4), 768–783.

Sharpe, W. F. 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance, 19(3), 425–442.