5.1 Introduction

Any increase in our understanding of how exchange rates are determined that improves our ability to predict exchange rate movements reduces exchange rate risk and, therefore, promotes more international economic activity. This handout presents the three most elementary and basic theories of exchange rate determination. The fact that there are three (and even others not presented in this handout) competing theories all trying to explain the same phenomenon is a telltale sign of our limited knowledge of exchange rates. But each of the three theories presented below has particular strengths that provide valuable insights to exchange rate behavior, even though they may not (yet) be woven together as part of a unified theory of exchange rates. The three theories are known as: purchasing power parity, the monetary approach to exchange rates, and the portfolio balance approach (a.k.a., asset market approach) to exchange rates.

5.2 Purchasing Power Parity (PPP)

Purchasing power parity (PPP) is aptly named because it maintains that there should be parity in the purchasing power of a given amount of money, regardless of which currency it is in or can be exchanged for. If the price of a widget in the US is $10 (i.e., the purchasing power of $10 is one widget), then that same $10, whether exchanged for ¥, £’s, or €’s, should be just enough to buy a comparable widget in Japan, England, and Europe, respectively. For example, if $\frac{¥}{£} = 100$ when the price in the US is $10, then PPP would predict the price in Japan to be ¥1000. Put more succinctly, PPP predicts that

\[ (P_{\text{DOM}}(e)) = P_{\text{FOR}} \]  

where \( P_{\text{DOM}} \) and \( P_{\text{FOR}} \) are the domestic and foreign prices. Or

\[ e_{\text{PPP}} = \frac{P_{\text{FOR}}}{P_{\text{DOM}}} \]  

where \( e_{\text{PPP}} \) is the exchange rate predicted by PPP. Equation 5.2 expresses PPP theory in a nutshell by highlighting its central prediction that the exchange rate will be the ratio of the foreign to domestic prices. Moreover, PPP theory provides a very good reason why PPP ought to hold. Once again, the force that promotes parity is arbitrage, and demonstrating how it maintains PPP is effectively accomplished by considering what happens when PPP does not hold.

Imagine that the price of a widget in the US is $10 (\( P_{\text{US}} = $10 \)), \( e_{\text{¥/£}} = 100 \), and the price in Japan for the widget (\( P_{\text{JAPAN}} \)) is ¥900. In this case PPP does not hold as \( (P_{\text{US}})(e_{\text{¥/£}}) > P_{\text{JAPAN}} \) and, therefore, there is an arbitrage opportunity (assuming transaction costs are not sufficiently large). Because \( P_{\text{JAPAN}} \) is on the smaller side of the inequality, one would want to buy the widget in Japan (for the ¥900), take the good and sell it in the US (for the $10), and then visit the foreign exchange market where, given \( e_{\text{¥/£}} \), the $10 buys ¥1000. Thus, the arbitrage converts the initial ¥900 to ¥1000. But the act of arbitrage always pushes prices to levels where arbitrage is no longer possible, which, in this case, is where PPP holds: All the arbitrageurs buying widgets in Japan would push up \( P_{\text{JAPAN}} \), the increase supply of widgets in the US would cause \( P_{\text{US}} \) to fall, and the increased supply of $ in the foreign exchange market would cause \( e_{\text{¥/£}} \) to fall, all three of which would tend to bring about PPP, given that \( (P_{\text{US}})(e_{\text{¥/£}}) > P_{\text{JAPAN}} \) to begin with.

As with the other parities already discussed, the degree to which arbitrage imposes PPP is a function of transaction costs. Continuing with the example above where \( P_{\text{US}} = $10 \), \( P_{\text{JAPAN}} = ¥900 \), and \( e = 100 \), if the act of transporting the widget from Japan to the US cost $1 or more, then no profit from arbitrage would be possible. More generally, if

\[ |(P_{\text{DOM}}(e)) - P_{\text{FOR}}| \leq \text{transaction costs} \]  

(5.3)
then arbitrage will not take place.

The largest transaction cost is often the cost of transporting the good from one market to another, especially in the case where the good is large, heavy, and/or perishable. But other transaction costs include the retailer’s costs and, in some cases, a tariff levied by the government of the country where the good is sold. One complicating factor is that some goods in one country are actually different than any goods sold in another. For example, T.V.’s in Europe have different specifications including running on different electrical current (DC instead of AC) such that a European T.V. cannot be used in the U.S.

It is true that estimates of $e_{PPP}$ rarely match actual values of $e$, but there seems to be limits to how different the two can be. There is a consensus that arbitrage will keep them from diverging too far apart, and that any observed deviation between the two is due to the transaction costs. It has become common to compare them and label the domestic currency as “overvalued” when $e > e_{PPP}$ and “undervalued” when $e < e_{PPP}$. Some even infer that an “overvalued” currency is destined to depreciate over time, and an “undervalued” currency will appreciate.

Although estimating $e_{PPP}$ is complicated in that the prices of many goods have to be entered into the calculation, it is possible to simply take existing measures of inflation in two countries and use PPP to predict changes in $e$ over that period. Rewriting Equation 5.2 by taking the natural logs of both sides and differencing gives:

$$\%\Delta e_{PPP} = %\Delta P_{FOR} - %\Delta P_{DOM} = \pi_{FOR} - \pi_{US}$$

(5.4)

For example, say that the inflation in the domestic country ($\pi_{DOM}$) was 12% while the foreign country’s inflation ($\pi_{FOR}$) = 8%. If PPP initially held (i.e., if $e_{PPP} = e$ before the inflation), then PPP would predict that $e$ would depreciate by 4% due to the inflation. But, of course, if the domestic currency was initially “overvalued” or “undervalued”, then PPP’s prediction of what happens to $e$ would have to be adjusted accordingly.

One problem with PPP as a theory is that, because arbitraging goods such as widgets is more involved and requires more time than arbitraging currencies or other paper assets, PPP only explains slow and deliberate movements in the exchange rate. It does not explain the variations observed in $e$ during the course of a day, week, or even months. An associated and conspicuously troubling aspect of PPP is that it completely overlooks KFA activity. As will be shown below, it is likely that much of the short run variation in $e$ that is not explained by PPP is due to KFA activity.

5.3 The Monetary Approach to the Exchange Rate

Monetarist Theory (or, Monetarism) maintains that changes in the aggregate supply and/or demand for money is a root cause of macroeconomic phenomena. The monetary approach to the exchange rate (MA) is a theory explaining exchange rate behavior that is guided by monetarist theory. Understanding changes in the aggregate money supply offers little mystery or challenge. The money supply ($M^S$) is the product of the monetary base and the “money multiplier” (a.k.a., demand deposit multiplier), and the monetary authority determines the exact amount of the monetary base and has great control over the money multiplier. Accordingly, the monetary authority has a firm grip on the country’s money supply and is usually held responsible for any changes in that money supply. The purchasing power of the money supply can be measured by dividing $M^S$ by the price level ($P$) to form what is know as real balances (or $\frac{M^S}{P}$).

In contrast, the aggregate demand for money in an economy, or a country’s money demand, is an interesting and complicated subject. Monetarist theory maintains that the
people’s desire to hold a particular amount of money either is, or reflects, the central behavior that characterizes a macroeconomy. One can think of an individual’s money demand at a given moment as the amount of money he or she wants to hold. Someone who has $25 in his or her pocket (and bank account) that is content not to spend any of it evidently has a money demand of $25. Someone who has $25, but purchases a good or asset for $10 reveals that his or her money demand is now $15. The theory maintains that people demand money in real terms, since it is the purchasing power of the money that they care about. Therefore, money demand is often defined as the amount of real balances that people want to hold in the economy and is represented by the symbol L.

Money demand is often thought of as the sum of two component demands: The transactions demand for money and the asset demand for money. There is a transactions demand for money (L\textsubscript{T}) because it is needed and used in most transactions. Even if purchases are made using credit, money is needed at some point to pay off the balance. The amount of real balances that people want to hold is a positive function of income (Y) and wealth (W): increases in either of these variables causes people to want to hold more real balances for transactions reasons.

It has also been made evident that people demand real balances beyond the amount desired for transactions purposes. There are several explanations why these additional balances are held, but all of them rely on money being an asset, thus, the real balances demanded beyond those desired for transactions is called the asset demand for money (L\textsubscript{A}). Money is often called an “inferior” asset because it earns no nominal return, but its role as the unit of account is clearly a valuable attribute. As the unit of account, its price cannot drop in the way the price of stocks, bonds, and other assets can. Thus, it is often a desirable asset in peoples’ portfolios when they are concerned about falling asset prices. Additionally, its liquidity as the unit of account makes it a popular asset to hold in the event of an emergency that requires sudden spending. These “precautionary balances” contribute to L\textsubscript{A}.

Other noteworthy determinants of L\textsubscript{A} are the risks associated with nonmonetary assets (e.g., increased uncertainty about the future value of equities makes money relatively more attractive) and changes in tastes (e.g., the allocation of wealth between the different nonmonetary assets and money is, to some degree, simply due to personal preferences or tastes). But regardless of the motive for L\textsubscript{A}, the total amount varies with the return earned on other assets such as the interest rate (i) earned on bonds and the yield (y) earned on equities. An increase in such returns would increase the opportunity cost of holding money (since money earns no comparable return), causing L\textsubscript{A} to fall. Similarly, a decrease in the returns on nonmonetary assets makes money look less costly to hold, and L\textsubscript{A} rises. Generally speaking, L\textsubscript{A} is a negative function of i. The transactions and asset demands for money add up to the entire money demand:

\[ L = L\textsubscript{T} + L\textsubscript{A} \] \hspace{1cm} (5.5)

Monetarism and the monetarist approach to the exchange rate maintains that equilibrium in the money market exists when

\[ (L = M\textsuperscript{S}/P) \] \hspace{1cm} (5.6)

and indirectly indicates that all the other major markets are in equilibrium, including the foreign exchange markets. It is only when \( L \neq M\textsuperscript{S}/P \) that notable events occur. For example, assume that L becomes less than \( M\textsuperscript{S}/P \) for some reason. This can be thought of as people having more money in their wallets, purses, and banks than they want. People will react to this disparity by unloading the unwanted real balances and purchasing one or both of the only possible options:
goods (and/or services) and non-money (i.e., “nonmonetary”) assets. Regardless of the combination of the goods and nonmonetary assets they buy, some of them will almost certainly be foreign, which necessarily includes selling the domestic currency in the foreign exchange markets – i.e., increasing the supply of domestic currency – that pushes e down. Therefore, the monetarist approach to exchange rates claims that an excess of real balances over money demand leads to a depreciation of the currency.

Monetarist theory also holds that such an excess of real balances over money demand is not an equilibrium situation. Of course, the spending of the money just redistributes the money supply to different holders and does not actually change its level – i.e., the unwanted money may circulate as if in a large scale game of hot potato – but the activity does precipitate other changes that cause L to equal M^S/P. The increased spending on goods will either stimulate an economy and increase its output (Y) and/or bring on inflation increasing its P. The increase in Y will cause L_T to rise, while the increase in P will cause M^S/P to fall, both of which push the initial inequality (i.e., L < M^S/P) towards equality. In addition, to the extent that the unwanted real balances are spent on nonmonetary assets, their prices will be bid up. An increase in the price of bonds constitutes a reduction in interest rates (KNOW WHY THIS IS TRUE) just as an increase in the price of equities reflects a decrease in y, both of which cause L_A to rise until L = M^S/P once again. Therefore, any excess in real balances will cause e to fall until the resulting rise in L_A and/or P restores L = M^S/P and everything, including e, settles at their new equilibrium values. The story when L > M^S/P is essentially symmetric: The excess money demand causes spending to slow or stop as people try to accumulate money in their wallets, purses, and banks. Some of the spending no longer occurring is the purchases of imported goods and, thus, the supply of domestic currency in the foreign exchange markets falls and e rises. The other tactic used to increase money balances is to sell nonmonetary assets, some of which are likely foreign. Purchasing the domestic currency in the foreign exchange markets with the proceeds from those foreign asset sales will also drive e up. Meanwhile, the reduction in spending activity also lowers domestic Y and P, and the sale of domestic nonmonetary assets will push up interest rates and yields, all causing L and M^S/P to seek the same level. Thus, an excess money demand will cause e to rise to a new equilibrium level where L = M^S/P.

The Monetarist approach may initially appear to have an unrealistic and bizarre focus on people’s money demand. For example, the approach characterizes people’s spending and investment decisions as derivative of their desire to hold more or less money. And who thinks of a person’s decision to purchase an apple the result of a decrease in money demand? But, upon serious thought, it is a legitimate way to look at things: If someone who previously preferred to hold money instead of buying an apple, now chooses to buy the apple, one can either say that the person’s desire for the apple relative to money rose, or, the desire for money relative to the apple fell. More importantly, the monetary approach accurately predicts how many changes or shocks in an economy effect e.

Employing the monetary approach to assess the effects of different kinds of shocks on e requires that one first identify how a shock affects either real balances or money demand. For example:

1) An increase in M^S – would cause real balances to exceed money demand which, given the monetarist story just presented, would cause e to fall.

2) An increase in P – would cause money demand to exceed real balances, which, given the monetarist story, would cause e to rise.
3) An increase in Y or W – would cause money demand to rise above real balances, which the monetarist approach would predict would cause e to rise.

The Monetarist approach does suffer a serious flaw that will be identified and examined after presenting the third exchange rate theory: the Portfolio Balance approach to the exchange rate.

5.4 The Portfolio Balance Approach (a.k.a., Asset Market Approach) to the Exchange Rate

Before formally presenting the portfolio balance approach to the exchange rate, a simple exercise is presented to provide helpful insight to the basic workings of asset markets and their derivation of asset prices. Assume that there are only two kinds of assets: A bonds and B bonds. At a given moment, there is a fixed number of each kind of bond and, therefore, the moment’s market diagram for each bond has a vertical supply (see Figure 5.1 and Figure 5.2). The demand for each asset is in part a function of the return in the other market. If A and B bonds are identical in every way (i.e., perfect substitutes for each other), then anyone (familiar with the concept of interest parity) would expect each bond’s demand to intersect the respective supply at the same common price. Now consider some possible changes:

1) Assume that A bonds become riskier than B bonds, so that now the average investor has to be compensated by a risk premium 5% on A bonds than B bonds. The demand for the riskier A bonds falls and the demand for B bonds rises as people try to shift some of their assets from the riskier to safer investments (see Figures 5.1 and 5.2). The falling demand for A bonds reduces their price, which corresponds to an increase in the rate on their rate of return. Similarly, the increasing demand for B bonds raises their price and lowers their return. This shift continues until the return on A bonds includes the 5% premium above the return on B bonds that fully compensates investors for the greater risk of holding A bonds. One the price of A bonds has fallen and B bonds has risen sufficiently to reflect the 5% risk premium, the investors stop redistributing their wealth from A to B bonds and the asset markets are once again in equilibrium.

2) Now resume the assumption that A and B bonds are perfect substitutes and consider the effect on both markets of a new issue of A bonds (see Figure 5.3). The increased supply of A bonds will lower the price of A bonds, thus increasing the interest rate on A bonds. Those who were previously happy holding B bonds as perfect substitutes for A bonds, will now want to sell their B bonds to purchase the higher returning A bonds. I.E., the demand for B bonds will fall and the demand for A bonds will rise. The decreased demand for B bonds will cause their price to fall and return to rise, the increased demand for A bonds will cause their price to rise and return to fall. Thus, the shift in demands in response to the higher interest rate on A bonds causes the return on A bonds and B bonds to converge (as shown in Figure 5.3).

This simple two-asset framework illustrates how a change that affects one asset will also affect other assets. The first example above shows how an increase in demand for one kind of asset cannot be separated from the associated decrease in demand for some other asset. The second example shows how a change in the supply of an asset will precipitate a desire by individuals to modify the portfolio of assets that they hold. Then, as in the first example, the acquisition of one type of asset is financed by the sale of another.¹ The portfolio balance approach to the exchange

¹ A similar mental exercise in which bonds and money are assumed to be the only two assets effectively captures a fundamental intuition behind the relationship between money and interest rates that was part of the monetary approach story related above. I.E., a change in preferences that motivates people to alter their portfolio to include
rates simply expands this basic interrelationship between different assets to account for the case when they (e.g., A bonds and B bonds) are denominated in different currencies.

The simple version of the portfolio balance approach assumes that there are only four kinds of assets in the world: domestic money (DM), domestic nonmonetary assets (DNMA), foreign nonmonetary assets (FNMA), and foreign money (FM). This situation differs from the two-asset example because it recognizes that the prices of two nonmonetary assets (and the corresponding rates of return) are also influenced by the respective money supplies as they were under the monetarist approach. For example, an increase in money supply that increases real balances above money demand will lead to more spending on assets (as well as goods) that will bid up the price of assets and lower their returns. But even with this extension of the model, the relative prices of the domestic and foreign assets are interconnected in the same way as A bonds and B bonds were in the example. Therefore, given the existing quantities of the four kinds of assets and the preferences for them by potential asset holders, there exists prices for the domestic and foreign nonmonetary assets that will generate a quantity demanded for each of the four assets that equals the existing quantity of that asset (in the same way that prices cleared both the Bond A and Bond B markets in the example above), and the financial markets will be in equilibrium.

The portfolio balance story begins with the prices of DNMA and FNMA such that the quantity demanded for each equals their existing supplies, i.e.,

\[
\begin{align*}
D_{DM} &= S_{DM} \\
D_{DNMA} &= S_{DNMA} \\
D_{FNMA} &= S_{FNMA} \\
D_{FM} &= S_{FM}
\end{align*}
\]

as every asset owner is content with the mix of nonmonetary assets and currencies they hold. The next step is to consider a change that causes people to shift their wealth to a different mix of these assets. Of course, there are two different kinds of changes that can occur: changes in either the supply or demand for one of these four assets.

Reasons for changes in any of the supplies is pretty straightforward. The existing amount of each currency is under the charge of the respective monetary authorities, while the nonmonetary assets are issued by both private entities in each country as well as both governments by issuing new bonds. The supplies of the four assets are usually thought of as fixed at any given moment and are vertical in asset market diagrams (as they were in Figures 5.1 – 5.3), with changes represented by a lateral shift of the corresponding supply curve.

Just as a change in the riskiness of A bonds in the above example that directly affected the demand for A bonds and, in so doing, indirectly affected the demand for B bonds, any variable affecting the demand for any of these four types of assets also influences at least one other, and usually all three of the other demands. The most important variables influencing the demands for the assets are: the domestic income or wealth (Y_{DOM} or W_{DOM}), real interest rate (r_{DOM}), expected inflation (\pi_{DOM}^{E}), and prices (P_{DOM}), the same variables for the foreign country

---

more bonds constitutes an increase in bond demand that is matched by a decrease in the demand for other assets, which in this example must be a decrease in money demand. The attempts to purchase the bonds with the money will drive up bond prices until the interest rate drops sufficiently to cause money demand to return to its former level. The effect of an increase in money supply can be assessed using similar reasoning: New money balances introduced when people were satisfied with the mix of bonds and money in their portfolios would likely make their portfolios top heavy in money. The selling of the unwanted money to buy bonds would raise bond prices, thus lowering interest rates until money demand rose enough to equal the increased money balances.
(Y_{FOR} or W_{FOR}, r_{FOR}, \pi_{E, \FOR}' \text{, } p_{FOR}), \text{ the relative risk premium } (rrp \text{ – defined in Handout #4), and the expected appreciation of the domestic currency } (xa \text{ – also defined in Handout #4). The variables directly responsible for each of the respective demands are:}

\begin{align*}
D_{DM} &= f_1(W_{DOM}, r_{DOM}, \pi_{E, DOM}', p_{DOM}, r_{FOR}, \pi_{E, FOR}', rrp, xa) \quad (5.7) \\
D_{DNMA} &= f_2(W_{DOM}, r_{DOM}, \pi_{E, DOM}', p_{DOM}, W_{FOR}, r_{FOR}, \pi_{E, FOR}', p_{FOR}, rrp, xa) \quad (5.8) \\
D_{FNMA} &= f_3(W_{DOM}, r_{DOM}, \pi_{E, DOM}', p_{DOM}, W_{FOR}, r_{FOR}, \pi_{E, FOR}', p_{FOR}, rrp, xa) \quad (5.9) \\
D_{FM} &= f_4(r_{DOM}, \pi_{E, DOM}', W_{FOR}, r_{FOR}, \pi_{E, FOR}', p_{FOR}, rrp, xa) \quad (5.10)
\end{align*}

Where the “+” or “-“ symbol above the variable indicates whether that variable has a positive or negative influence on that particular demand. For example, the “+” above \( \pi_{E, FOR}' \) in the \( D_{DM} \) equation indicates that an increase in \( \pi_{E, FOR}' \) will increase \( D_{DM} \), while a decrease in \( \pi_{E, FOR}' \) will decrease \( D_{DM} \). As seen in Equation 5.8, the change in \( \pi_{E, FOR}' \) will also cause the \( D_{DNMA} \) to move with it, while the “-“ above \( \pi_{E, FOR}' \) in Equation 5.9 indicates that \( D_{FNMA} \) moves in the opposite direction from changes in \( \pi_{E, FOR}' \). As another example, an increase in \( r_{DOM} \) will cause \( D_{DM} \) to fall, \( D_{DNMA} \) to rise, and \( D_{FNMA} \) to fall since it makes DNMA relative more attractive to now hold.

Knowing how the supplies and demands for the four assets are determined now allows us to implement the portfolio balance approach. Consider the following cases:

1) An increase in the supply of FNMA in the form of newly issued bonds (in the same way new A bonds were issued in Figure 5.3) – will cause the price of FNMA to fall which corresponds to an increase in \( f_{FOR} \) (FOR REASONS THAT ARE PERFECTLY CLEAR TO YOU, IF NOT…..). The increase in \( f_{FOR} \) causes \( D_{DM} \) to fall, \( D_{DNMA} \) to fall, and \( D_{FM} \) to fall as the world’s wealth holders shift their portfolios to hold more of the now more attractive FNMA. Because this reallocation involves people selling domestic assets to buy the FNMA, activity occurs in the foreign exchange market. In this case the supply of domestic currency would rise, causing e to fall. Therefore, the portfolio balance approach predicts that an increase in the supply of FNMA will cause e to fall.

2) An increase in \( xa \) – will cause \( D_{DM} \) to rise, \( D_{DNMA} \) to rise, \( D_{FNMA} \) to fall, and \( D_{FM} \) to fall. Investors will be moving from foreign assets to domestic, requiring the sale of foreign currency in the foreign exchange markets in return for domestic currency with which to purchase the domestic assets: e will rise. Therefore, the portfolio balance approach predicts that an increase in \( xa \) will cause e to rise.

Note that the portfolio balance approach is completely consistent with the principle of interest parity. For example, when the supply FNMA increases causing their prices to fall and their returns to rise, the decreased demand for DNMA caused their respective prices to fall and returns to rise as well.
Also note that although $W_{FOR}$ and $P_{FOR}$ do not directly affect $D_{DM}$ (see Equation 5.7) just as $W_{DOM}$ and $P_{DOM}$ do not directly influence $D_{FM}$ (see Equation 5.10), these variables will still indirectly affect those markets through their effect on the other three asset markets.

The monetary approach flaw: A strict application of the monetary approach to an increase in interest rates would predict that the decrease in money demand would cause spending on domestic and foreign goods and assets that would cause $e$ to fall. But it is generally agreed (with great confidence) that this is plainly wrong: Higher real interest rates attract foreign investment that dependably cause $e$ to rise. The problem lies in the fact that the monetary approach implicitly assumes that both domestic and foreign agents are (somehow) oblivious to the possibility that foreign interest rates might not be matching the change in the domestic rate. Therefore, the higher domestic rate has no direct effect on capital and financial account activity, but just the effects associated with agents attaining the desired level of real balances. The portfolio balance approach clearly does account for the direct effect of different interest rates on capital flows and predicts the effects of changes in real interest rates on $e$ correctly.

5.5 The Monetary and Portfolio Balance Approaches to the BOP

The monetary and portfolio balance approaches predictions regarding $e$ described above will obviously not occur when a government has a fixed exchange rate policy in place. If they are honoring a pegged rate, then any pressure for $e$ to fall will cause the government to purchase the domestic currency with reserves out of its vault, causing the $KFA_{GOV}$ to be in surplus and the BOP to be in deficit. Thus, a prediction by either approach that $e$ is to fall will, under the circumstances of a fixed $e$, be a prediction that BOP deficits will occur. Similarly, events that either theory would view as a cause for a flexible currency to appreciate would be predicted to generate BOP surpluses under a fixed rate policy.

Key Terms:

- asset demand for money
- Monetarist theory (or, Monetarism)
- monetary approach
- money demand
- portfolio balance approach (a.k.a., asset market approach)
- purchasing power parity (PPP)
- real balances
- transactions demand for money