Using Higher Order Moments to Estimate Equations with Differential Measurement Error

Karim Chalak*
University of Montreal

Daniel Kim
BI Norwegian Business School

July 30, 2022

Abstract

We study a leading case of differential measurement error that occurs when a proxy for a latent variable directly affects the response of interest, thereby violating the proxy exclusion restriction. We show that using higher order moments, together with reasonable sign restrictions, can point identify all the equation coefficients.

PRELIMINARY AND INCOMPLETE DRAFT

JEL codes: C21, C26.

Keywords: differential measurement error, exclusion restriction, higher order moments, independence.

*Karim Chalak, Department of Economics, University of Montreal, karim.chalak@umontreal.ca. Daniel Kim, BI Norwegian Business School, daniel.s.kim@bi.no.

1 Introduction

Researchers frequently confront the challenge that data on variables of interest U are not perfectly available. Instead, one may have access to an error-laden proxy W for the latent variables U. For example, when estimating the effects of the quality of a product (e.g. an asset or a movie) on its demand, one may only have access to the product's rating (e.g. asset rating or movie review). It is well known that simply substituting an error-laden proxy W for U results in biased estimates of the effects of U on the outcome Y. Instead, one should appropriately account for the noise in the data. In particular, if the outcome equation for Y is linear and the measurement error is classical, then one can bound the effect of U on Y (see e.g. Gini, 1921; Frisch, 1934; Klepper and Leamer, 1984; Bollinger, 2003; Chalak and Kim, 2020). Moreover, if the measurement error, latent variable, and equation disturbance are jointly independent, as opposed to merely uncorrelated as in the classical case, then one can use the moments of order higher than two to point identify the equation coefficients without resort to extraneous information (see e.g. Reiersøl, 1950; Dagenais and Dagenais, 1997; Erickson and Whited, 2002; Erickson, Jiang, and Whited, 2014).

These standard partial and point identification results hold under the assumption that the proxy is excluded from the outcome equation. However, a researcher may wish to relax this exclusion restriction (see e.g. the discussion in Chalak and Kim, 2021). For instance, a product's rating (e.g. asset rating or movie review) that serves as a proxy for the product's quality in the researcher's analysis, may directly influence the product's demand. Similarly, the result of a medical test may influence a patient's behavior (e.g. a worker may work less hours if she is incorrectly prescribed rest). In this case, the measurement error is "differential," since the distribution of $Y|(U,W)$ differs from that of $Y|U$, and the standard results are not directly applicable. Indeed, Chalak and Kim (2021) demonstrate that, relaxing the proxy exclusion restriction invalidates the classical bounds, and the equation coefficients are no longer identified (i.e. the identification regions span the real line).

This paper puts forward constructive identification results in the relatively less explored, yet more general, setting of differential measurement error (see e.g. the discussion in Bound, Brown, and Mathiowetz, 2001). In particular, the paper relaxes the ubiquitous proxy exclusion restriction to allow for nonzero coefficients on the proxy variables in the outcome equation. In this leading setting for differential measurement "W is not merely a mismeasured version of $[U]$, but is a separate variable acting as a type of proxy for $[U]$" (Carroll, Ruppert, Stefanski, and Crainiceanu, 2006, p. 36). The paper then demonstrates that higher order moments can be used to point-identify the effects of latent variables U, as well as of their proxies W, on the outcome Y. In the course of our analysis, we also relax some of assumptions that are sometimes maintained in the literature on nondifferential measurement error. Specifically, we relax the assumption that the measurement errors are jointly independent (see e.g. Erickson and Whited (2002) and Erickson, Jiang, and Whited (2014)) to allow for arbitrary dependence. Nor do we impose distributional assumptions, such as requiring that the disturbance and measurement errors are normally distributed (see e.g. Dagenais and Dagenais, 1997).

The rest of the paper is organized as follows. Section 2 introduces the assumptions. Section 3 studies the identification of the equation coefficients and other quantities and comments on special cases of interest. Section 4 studies estimation and inference. Section 5 presents simulation results. Section 6 concludes.

2 Data Generation and Assumptions

We consider the following data generating assumptions.

Assumption $\mathbf{A_1}$ *Linearity: (i) Let the random vector* $(\underset{J\times 1}{W}', \underset{1\times 1}{Y})'$, *with* $J \geq 1$, *have sufficiently many finite moments. (ii) Let the random variables* $\underset{1\times 1}{\eta}, \underset{J\times 1}{U}, \underset{J\times 1}{\varepsilon}$, W, *and* Y *satisfy*

$$Y = W'\phi + U'\delta + \eta \quad \text{and} \quad W = U + \varepsilon \tag{1}$$

with constant slope coefficients. Set $E(\eta) = 0$, $E(U) = 0$, *and* $E(\varepsilon) = 0$. *The researcher observes realizations of* $(W', Y)'$ *but not of* (U', η, ε').

A_1 specifies the outcome equation and decomposes the proxy W into the latent variables (signal) U and the measurement error (noise) ε. For simplicity, all the variables are demeaned. A_1 allows the coefficient ϕ on W to be nonzero. As such the distributions of $Y|(U, W)$ and $Y|U$ generally differ, and the measurement error is "differential." We note that A_1 assumes that all the explanatory variables U in the Y equation are potentially mismeasured. In section 3.3.3, we comment on the special case in which the researcher wishes to assume that a subset X of the explanatory variables is perfectly measured.

Assumption $\mathbf{A_2}$ *Joint Independence:* $U \perp (\eta, \varepsilon')$ *and* $\eta \perp \varepsilon$.

A_2 assumes that U, η, and ε are jointly independent. Note that A_2 does not require the components of ε to be uncorrelated or independent. Nor does it restrict the dependence among the residuals, so $Var(\eta)$ need not be diagonal.

3 Identification

3.1 A Scalar Latent Variable

It is instructive to begin by examining the case in which $J = 1$, with U and W scalars. Together, A_1 and A_2 allow expressing the moments involving (W, Y) as a function of a vector θ^* of unknown parameters. For example, using the moments of order at most 4, we can generate a system involving the following 9 moment equations:

$$\begin{aligned}
\mu_{W,Y} &= (\phi+\delta)\mu_U^2 + \phi\mu_\varepsilon^2 \qquad (2)\\
\mu_W^2 &= \mu_U^2 + \mu_\varepsilon^2\\
\mu_Y^2 &= (\phi+\delta)^2\mu_U^2 + \phi^2\mu_\varepsilon^2 + \mu_\eta^2\\
\mu_{W,Y}^{2,1} &= (\phi+\delta)\mu_U^3 + \phi\mu_\varepsilon^3\\
\mu_{W,Y}^{1,2} &= (\phi+\delta)^2\mu_U^3 + \phi^2\mu_\varepsilon^3\\
\mu_W^3 &= \mu_U^3 + \mu_\varepsilon^3\\
\mu_{W,Y}^{3,1} &= (\phi+\delta)\mu_U^4 + \phi\mu_\varepsilon^4 + 3(\phi+\delta)\mu_U^2\mu_\varepsilon^2 + 3\phi\mu_U^2\mu_\varepsilon^2\\
\mu_{W,Y}^{2,2} &= (\phi+\delta)^2\mu_U^4 + \phi^2\mu_U^2\mu_\varepsilon^2 + \mu_U^2\mu_\eta^2 + 4(\phi+\delta)\phi\mu_U^2\mu_\varepsilon^2 + (\phi+\delta)^2\mu_U^2\mu_\varepsilon^2 + \phi^2\mu_\varepsilon^4 + \mu_\varepsilon^2\mu_\eta^2, \text{ and}\\
\mu_{W,Y}^{1,3} &= (\phi+\delta)^3\mu_U^4 + 3(\phi+\delta)\mu_U^2\mu_\eta^2 + 3(\phi+\delta)\phi^2\mu_U^2\mu_\varepsilon^2 + \phi^3\mu_\varepsilon^4 + 3(\phi+\delta)^2\phi\mu_U^2\mu_\varepsilon^2 + 3\phi\mu_\varepsilon^2\mu_\eta^2,
\end{aligned}$$

and the following 9 unknowns[1] appearing on the left hand side:

$$\theta^* \equiv (\phi, \delta, \mu_U^2,\ \mu_\varepsilon^2,\ \mu_\eta^2,\ \mu_U^3,\ \mu_\varepsilon^3,\ \mu_U^4,\ \mu_\varepsilon^4).$$

Note that the moment $\mu_Y^3 = (\phi+\delta)^3\mu_U^3 + \phi^3\mu_\varepsilon^3 + \mu_\eta^3$ involves the unknown μ_η^3 which does not appear in the above equations. If μ_η^3 is of direct interest, adding μ_Y^3 results in a larger system of 10 equations and 10 unknowns. A similar comment applies to μ_Y^4.

[1] Recall that by A_1, ε and η each have zero mean.

As is well known when $\phi = 0$ and the measurement error is nondifferential, here too, there are instances in which the above system of equations has infinitely many roots. In this case, θ^* is not point-identified. For example, suppose that the distribution of U is symmetric, $\mu_U^3 = 0$, and that either ε has a symmetric distribution, $\mu_\varepsilon^3 = 0$, or $\phi = 0$ (or both). Then the moments of order three ($\mu_{W,Y}^{2,1}$, $\mu_{W,Y}^{1,2}$, and μ_W^3) are uninformative about the other components of θ^*, and the system is underdetermined.

Beyond these cases, the next theorem gives general conditions under which the above system admits only two roots despite suffering from differential measurement error. Further, simply imposing a sign restriction on δ point identifies the system parameters θ^*. To derive this result, theorem 3.1 makes use of the moment equations (2) to express all the components of θ^* as a function $M(\cdot)$ of ϕ (as well as of estimable moments of (Y, W) which this notation leaves implicit). This reduces the problem of solving the system of equations (2) in θ^* to the problem of solving a nonlinear equation in ϕ. Theorem 3.1 shows that this equation admits two roots f_1 and f_2. Further, it shows that these two roots imply two values $d_1 = M_\delta(f_1)$ and $d_2 = M_\delta(f_2)$ for δ that are of equal magnitude but of opposite signs. It follows, that restricting the sign of δ identifies the unique root f^* for ϕ and therefore point identifies θ^*

via the mapping $M(f^*)$. The components of $M(\cdot)$ can be expressed recursively by:

$$
\begin{aligned}
\phi &= M_\phi(\phi) \equiv \phi, \\
\mu_\varepsilon^3 &= M_\varepsilon^3(\phi) \equiv \frac{\mu_W^3 \mu_{W,Y}^{1,2} - (\mu_{W,Y}^{2,1})^2}{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}, \\
\delta &= M_\delta(\phi) \equiv M_{\phi+\delta}(\phi) - \phi \equiv \frac{\mu_{W,Y}^{1,2} - \phi^2 M_\varepsilon^3(\phi)}{\mu_{W,Y}^{2,1} - \phi M_\varepsilon^3(\phi)} - \phi, \\
\mu_U^3 &= M_U^3(\phi) \equiv \frac{\mu_{W,Y}^{2,1} - \phi M_\varepsilon^3(\phi)}{M_{\phi+\delta}(\phi)}, \\
\mu_\varepsilon^2 &= M_\varepsilon^2(\phi) \equiv \frac{\mu_{W,Y} - M_{\phi+\delta}(\phi)\mu_W^2}{-M_\delta(\phi)}, \\
\mu_U^2 &= M_U^2(\phi) \equiv \mu_W^2 - M_\varepsilon^2(\phi), \\
\mu_\eta^2 &= M_\eta^2(\phi) \equiv \mu_Y^2 - M_{\phi+\delta}(\phi)^2 M_U^2(\phi) - \phi^2 M_\varepsilon^2(\phi), \\
\mu_\varepsilon^4 &= M_\varepsilon^4(\phi) \\
&\equiv \frac{\mu_{W,Y}^{2,2} - M_{\phi+\delta}(\phi)\mu_{W,Y}^{3,1} + M_\delta(\phi)(\phi + 2M_{\phi+\delta}(\phi))M_U^2(\phi)M_\varepsilon^2(\phi) - M_U^2(\phi)M_\eta^2(\phi) - M_\varepsilon^2(\phi)M_\eta^2(\phi)}{-\phi M_\delta(\phi)}, \\
\mu_U^4 &= \frac{\mu_{W,Y}^{3,1} - \phi M_\varepsilon^4(\phi) - 3(M_{\phi+\delta}(\phi) + \phi)M_U^2(\phi)M_\varepsilon^2(\phi)}{M_{\phi+\delta}(\phi)}.
\end{aligned}
$$

Theorem 3.1 *Assume A_1 and A_2 and let ϕ, δ, and μ_U^3 be nonzero. Define*

$$
\begin{aligned}
A &\equiv \mu_{W,Y}^{2,1}(3\mu_{W,Y}\mu_W^2 - \mu_{W,Y}^{3,1}) + \mu_W^3(\mu_{W,Y}^{2,2} - 2\mu_{W,Y}\mu_{W,Y} - \mu_W^2\mu_Y^2), \\
B &\equiv \mu_W^3(-\mu_{W,Y}^{1,3} + 3\mu_{W,Y}\mu_Y^2) + \mu_{W,Y}^{1,2}(\mu_{W,Y}^{3,1} - 3\mu_W^2\mu_{W,Y}), \text{ and} \\
C &\equiv \mu_{W,Y}^{2,1}(\mu_{W,Y}^{1,3} - 3\mu_{W,Y}\mu_Y^2) - \mu_{W,Y}^{1,2}(\mu_{W,Y}^{2,2} - 2\mu_{W,Y}\mu_{W,Y} - \mu_W^2\mu_Y^2),
\end{aligned}
$$

and let $\Delta = B^2 - 4AC$.

(a) The system of equations (2) admits two roots for θ^* given by $M(f_1)$ and $M(f_2)$ where

$$
f_1 = \frac{-B - \sqrt{\Delta}}{2A.} \quad \text{and} \quad f_2 = \frac{-B + \sqrt{\Delta}}{2A.}.
$$

(b) Moreover, we have that the components of M corresponding to δ, U, and ε are such that $M_\delta(f_1) = -M_\delta(f_2)$, $M_U^2(f_1) = M_\varepsilon^2(f_2)$, and $M_U^2(f_2) = M_\varepsilon^2(f_1)$. Therefore, knowing either the sign of δ or/and that $\mu_\varepsilon^2 \le \mu_U^2$ would recover the correct root $f^* \in \{f_1, f_2\}$. Then $M(f^*)$ point identifies θ^*.

3.2 A Vector of Latent Variables

We generalize Section 2.1's analysis to accommodate a vector of latent variables, with $J \geq 1$. As in the scalar case, we express the higher order moments involving (W, Y) as a function of ϕ, δ, and of the higher order moments involving (U, ε, η). To shorten the notation, for non-negative integers $r_1, ..., r_d$, and distinct random variables $A_1, ..., A_d$, we denote the central moment of order $\sum_{h=1}^{d} r_h$ by:

$$\mu_{A_1,...,A_d}^{r_1,...,r_d} \equiv E[(A_1 - E(A_1))^{r_1} ... (A_d - E(A_d))^{r_d}].$$

By A_1, A_2, and the multinomial theorem, these moments take the following form, for $r = \sum_{j=1}^{J} r_j$ and $2 \leq r + q$:

$$\begin{aligned} \mu_{W_1,...,W_J,Y}^{r_1,...,r_J,q} &= E[(U_1 + \varepsilon_1)^{r_1} ... (U_J + \varepsilon_J)^{r_J} (U'(\phi + \delta) + \varepsilon'\phi + \eta)^q] \\ &= \sum_{d_1=0}^{r_1} .... \sum_{d_J=0}^{r_J} \sum_{c_1+...+c_{2J+1}=q} \frac{r_1!}{d_1!(r_1 - d_1)!} ... \frac{r_J!}{d_J!(r_J - d_J)!} \frac{q!}{c_1!...c_{2J+1}!} \\ & \prod_{s=1}^{J} (\phi_s + \delta_s)^{c_s} \;\; \mu_{U_1,...,U_J}^{c_1+r_1-d_1,...,c_J+r_J-d_J} \prod_{t=J+1}^{2J} \phi_{t-J}^{c_t} \;\; \mu_{\varepsilon_1,...,\varepsilon_J}^{c_{J+1}+d_1,...,c_{2J}+d_J} \;\; \mu_{\eta}^{c_{2J+1}}. \end{aligned} \tag{3}$$

The last sum is over all $(c_1, ..., c_{2J+1})$ combinations that sum to q. Setting $q = 0$ yields the moment $\mu_{W_1,...,W_J}^{r_1,...,r_J}$ whereas setting $r = 0$ yields the moment μ_Y^q.

By sequentially stacking the moments in increasing order, it is possible to generate a system with at least as many equations than unknowns. Specifically, recall that there are $\frac{(R+r-1)!}{r!(R-1)!}$ ways to select a combination of r items from R possibilities with repetition. Thus, ignoring the (zero) means where $l \equiv \sum_{j=1}^{J} l_j = 1$, there are $\sum_{l=2}^{r} \frac{(J+l)!}{l!J!}$ moments of the form $\mu_{W_1,...,W_J,Y}^{l_1,...,l_J,q}$ of order at most r. These moments involve the unknowns ϕ and δ, $2\sum_{l=2}^{r} \frac{(J+l-1)!}{l!(J-1)!}$ unknowns of the form $\mu_{U_1,...,U_J}^{l_1,...l_J}$ or $\mu_{\varepsilon_1,...,\varepsilon_J}^{l_1,...,l_J}$, and $r - 1$ unknowns of the form μ_η^l. In total, there are $\sum_{l=2}^{r} \frac{(J+l)!}{l!J!}$ moment equations involving $2J + 2\sum_{l=2}^{r} \frac{(J+l-1)!}{l!(J-1)!} + r - 1$ unknowns, which we collect in the vector θ^*. For a fixed J, the ratio of equations to unknowns diverges as r increases. For example, for $J = 2$, setting $r = 4$ yields a system of 31 moment equations involving 31 unknowns, whereas setting $r = 5$ yields a system of 52 equations and 44 unknowns. Similarly, setting $J = 3$ and $r = 5$ yields a system of 121 moment equation and 114 unknowns. As in the scalar case, certain moments (such as μ_Y^3 or μ_Y^4) may involve ancillary unknowns that do not appear elsewhere. If so, removing these moments reduces the dimension of the equations and unknowns. If the moment equations carry redundant information (as in the joint normality example discussed above), the system may have multiple, or infinitely many, solutions. Otherwise, θ^* may be unique solution to this system.

3.3 Key Special Cases and Discussion

We briefly comment on two special cases of interest that reduce the dimension of the unknowns $\tilde{\theta}^*$, and on the role of the perfectly measured covariates.

3.3.1 Multiple Proxies for a Latent Variable

Multiple proxies may be available for one or more latent variables. For example, suppose that the first m components of W are proxies for the same latent variable V_1. Then , for $j = 1, ..., m$, one can set $U_j = V_1$ and restrict the corresponding coefficients in the Y equation to $\delta_j = \frac{1}{m}\gamma_1$. This reduces the dimension of the unknowns in δ to $J - m + 1$ and reduces the number of unknowns of the form $\mu^{l_1,...l_H}_{U_{j_1},...,U_{j_H}}$ to $\sum_{l=2}^{r} \frac{(J-m+l)!}{l!(J-m)!}$.

3.3.2 Independent Measurement Errors

Our analysis allows the components of the measurement error vector ε to be correlated or dependent. The literature often assumes that the components of ε are jointly independent. Requiring $Var(\varepsilon)$ to be diagonal implies that $\mu^{l_1,...,l_H}_{\varepsilon_{j_1},...,\varepsilon_{j_H}} = \mu^{l_1}_{\varepsilon_{j_1}} ... \mu^{l_H}_{\varepsilon_{j_H}}$. Ignoring the means, this reduces the number of unknowns of order at most r that are of this form to $J \times (r - 1)$.

3.3.3 Perfectly Measured Covariates

If available, perfectly measured covariates X can be used to construct instruments (see e.g. Lewbel, 1997). As a simple example, let X enter the Y equation

$$Y = X'\beta + W'(\phi + \delta) - \varepsilon'\delta + \eta,$$

where we substitute for $U = W - \varepsilon$. If the covariates X are independent of (ε, η) then nonlinear functions of X may be used as instruments for (X, W) in the above equation to identify β and $\phi + \delta$. Note that this does not directly identify ϕ and δ separately. But one can consider supplementing the moments in Sections 3.1 and 3.2 with moments that involve

X. The analysis in Sections 3.1 and 3.2 does not require perfectly measured covariates - it is valid even when $X = 1$ and all the variables are measured with error.

4 Estimation and Inference

Let $Z = (W, Y)$ and recall that θ^* denotes the $p \times 1$ vector of unknowns, as in Section 3.1. Let the $p \times 1$ vector $m(Z)$ stack the moments of the form $\mu^{r_1,\ldots,r_H,q}_{W_{j_1},\ldots,W_{j_H},\tilde{Y}}$ that are generated by equation (3), and let $c(\theta^*)$ stack the corresponding left hand side functions of θ^*. This defines θ^* as a solution to the moment equations

$$E(\underset{p\times 1}{g(Z,\theta^*)}) = E(\underset{p\times 1}{m(Z) - c(\theta^*)}) = 0. \tag{4}$$

Given a square p positive definite weighting matrix Ξ^*, we have

$$\theta^* = \arg\min_{\theta\in\Theta} E[g(Z,\theta)]'\Xi^* E[g(Z,\theta)).$$

We can then estimate θ^* using an asymptotically normal and consistent generalized method of moments (GMM) estimator $\hat{\theta}$. The properties of GMM estimators are well-established - we omit a detailed discussion for brevity.

If covariates X are present then one must account for conditioning on these. In this case, the projections of Y and W on X would depend on unknown regression coefficients which satisfy certain regression moment equations. One can augment the above moments with these regression moment equations and define θ^* together with the regression coefficients as the solution to the stacked system of moments equations.

5 Simulation

6 Conclusion

References

Bollinger, C. (2003), "Measurement Error in Human Capital and the Black-White Wage Gap," *Review of Economics and Statistics*, 85, 578-585.

Bound, J., Brown, C., Mathiowetz, N. (2001), "Measurement error in survey data," in: Heckman, J. and Leamer, E. eds., *Handbook of Econometrics*, vol. 5.

Carroll, R., D. Ruppert, L. Stefanski, and C. Crainiceanu (2006). *Measurement Error in Nonlinear Models: A Modern Perspective.* Chapman and Hall/CRC, Second Edition.

Chalak, K. and D. Kim (2020), "Measurement Error in Multiple Equations: Tobin's q and Corporate Investment, Saving, and Debt," *Journal of Econometrics*, 214, 413-432.

Chalak, K. and D. Kim (2021), "Measurement Error without the Proxy Exclusion Restriction," *Journal of Business and Economic Statistics*, 39, 200-216.

Dagenais M. and D. Dagenais (1997), "Higher Moment Estimators for Linear Regression Models with Errors in the Variables," *Journal of Econometrics*, 76, 193-221.

Erickson, T. and T. Whited (2002), "Two-Step GMM Estimation of the Errors-in-Variables Model Using High-Order Moments," *Econometric Theory*, 18, 776-799.

Erickson, T., C. Jiang, and T. Whited (2014), "Minimum Distance Estimation of the Errors-in-Variables Model Using Linear Cumulant Equations," *Journal of Econometrics*, 183, 211-221.

Frisch, R. (1934). *Statistical Confluence Analysis by Means of Complete Regression Systems.* Oslo, Norway: University Institute of Economics.

Gini, C. (1921), "Sull'Interpolazione di una Retta Quando i Valori della Variabile Indipendente Sono Affetti da Errori Accidentali," *Metroeconomica*, 1, 63-82.

Klepper, S. and E. Leamer (1984), "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica*, 52, 163-184.

Lewbel, A. (1997), "Constructing Instruments for Regressions with Measurement Error when No Additional Data are Available, with an Application to Patents and R&D," *Econometrica*, 65, 1201-1213.

Reiersøl, O. (1950), "Identifiability of a Linear Relation between Variables Which Are Subject to Error," *Econometrica*, 18, 375-389.

7 Mathematical Proofs

Proof of Theorem 3.1: We use the system of equations (2) to express all the unknowns as a function of ϕ. We then solve for ϕ. First, using equations 4, 5, and 6 in (2), we have that

$$\begin{aligned} \mu_{W,Y}^{2,1} - \phi\mu_{\varepsilon}^{3} &= (\phi+\delta)\mu_{U}^{3} \\ \mu_{W,Y}^{1,2} - \phi^{2}\mu_{\varepsilon}^{3} &= (\phi+\delta)^{2}\mu_{U}^{3}. \end{aligned}$$

Given that $(\phi+\delta)$ and μ_U^3 are nonzero, we obtain

$$\begin{aligned}(\phi+\delta) &= \frac{\mu_{W,Y}^{1,2}-\phi^2\mu_\varepsilon^3}{\mu_{W,Y}^{2,1}-\phi\mu_\varepsilon^3}\\ \mu_U^3 &= \frac{\mu_{W,Y}^{2,1}-\phi\mu_\varepsilon^3}{(\phi+\delta)}\\ \mu_\varepsilon^3 &= \mu_W^3-\frac{\mu_{W,Y}^{2,1}-\phi\mu_\varepsilon^3}{(\phi+\delta)}.\end{aligned}$$

Combining the equations for $(\phi+\delta)$ and μ_ε^3 then gives

$$\mu_W^3-\mu_\varepsilon^3=(\mu_{W,Y}^{2,1}-\phi\mu_\varepsilon^3)\frac{\mu_{W,Y}^{2,1}-\phi\mu_\varepsilon^3}{\mu_{W,Y}^{1,2}-\phi^2\mu_\varepsilon^3}.$$

Rearranging this expression gives

$$\begin{aligned}&(\mu_W^3-\mu_\varepsilon^3)(\mu_{W,Y}^{1,2}-\phi^2\mu_\varepsilon^3)-(\mu_{W,Y}^{2,1}-\phi\mu_\varepsilon^3)^2\\ =\;&\mu_W^3\mu_{W,Y}^{1,2}-\phi^2\mu_W^3\mu_\varepsilon^3-\mu_\varepsilon^3\mu_{W,Y}^{1,2}+\phi^2(\mu_\varepsilon^3)^2-(\mu_{W,Y}^{2,1})^2+2\phi\mu_\varepsilon^3\mu_{W,Y}^{2,1}-\phi^2(\mu_\varepsilon^3)^2\\ =\;&\mu_W^3\mu_{W,Y}^{1,2}-\phi^2\mu_W^3\mu_\varepsilon^3-\mu_\varepsilon^3\mu_{W,Y}^{1,2}-(\mu_{W,Y}^{2,1})^2+2\phi\mu_\varepsilon^3\mu_{W,Y}^{2,1}=0.\end{aligned}$$

We can then express μ_ε^3, $(\phi+\delta)$, and μ_U^3 as functions of ϕ:

$$\begin{aligned}\mu_\varepsilon^3 &= M_\varepsilon^3(\phi)\equiv\frac{\mu_W^3\mu_{W,Y}^{1,2}-(\mu_{W,Y}^{2,1})^2}{\phi^2\mu_W^3+\mu_{W,Y}^{1,2}-2\phi\mu_{W,Y}^{2,1}}=\frac{\mu_W^3\mu_{W,Y}^{1,2}-(\mu_{W,Y}^{2,1})^2}{\phi(\phi\mu_W^3-\mu_{W,Y}^{2,1})-(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2})}\equiv\frac{S}{L(\phi)},\\ (\phi+\delta) &= M_{\phi+\delta}(\phi)\equiv\frac{\mu_{W,Y}^{1,2}-\phi^2M_\varepsilon^3(\phi)}{\mu_{W,Y}^{2,1}-\phi M_\varepsilon^3(\phi)},\text{ and}\\ \mu_U^3 &= M_U^3(\phi)\equiv\frac{\mu_{W,Y}^{2,1}-\phi M_\varepsilon^3(\phi)}{M_{\phi+\delta}(\phi)},\end{aligned}$$

where we require that $L(\phi)\neq 0$ Further, using equations 1, 2, and 3 in (2), we have

$$\mu_{W,Y}=(\phi+\delta)(\mu_W^2-\mu_\varepsilon^2)+\phi\mu_\varepsilon^2=M_{\phi+\delta}(\phi)\mu_W^2+(\phi-M_{\phi+\delta}(\phi))\mu_\varepsilon^2$$

and thus, we can express the variables in these equations as a function of ϕ. Given $\delta\neq 0$, we have

$$\begin{aligned}\mu_\varepsilon^2 &= M_\varepsilon^2(\phi)\equiv\frac{\mu_{W,Y}-M_{\phi+\delta}(\phi)\mu_W^2}{\phi-M_{\phi+\delta}(\phi)},\\ \mu_U^2 &= M_U^2(\phi)\equiv\mu_W^2-M_\varepsilon^2(\phi),\text{ and}\\ \mu_\eta^2 &= M_\eta^2(\phi)\equiv\mu_Y^2-M_{\phi+\delta}(\phi)^2M_U^2(\phi)-\phi^2M_\varepsilon^2(\phi).\end{aligned}$$

Last, equations 7, 8, and 9 in in (2) can be rewritten as

$$\begin{aligned}
\mu_{W,Y}^{3,1} &= (\phi+\delta)\mu_U^4 + \phi\mu_\varepsilon^4 + 3[(\phi+\delta)+\phi]\mu_U^2\mu_\varepsilon^2 \\
\mu_{W,Y}^{2,2} &= (\phi+\delta)^2\mu_U^4 + \phi^2\mu_\varepsilon^4 + [\phi^2+4(\phi+\delta)\phi+(\phi+\delta)^2]\mu_U^2\mu_\varepsilon^2 + \mu_U^2\mu_\eta^2 + \mu_\varepsilon^2\mu_\eta^2, \text{ and} \\
\mu_{W,Y}^{1,3} &= (\phi+\delta)^3\mu_U^4 + \phi^3\mu_\varepsilon^4 + 3(\phi+\delta)\mu_U^2\mu_\eta^2 + 3\phi(\phi+\delta)(\phi+(\phi+\delta))\mu_U^2\mu_\varepsilon^2 + 3\phi\mu_\varepsilon^2\mu_\eta^2.
\end{aligned}$$

Then

$$\begin{aligned}
\mu_U^4 &= \frac{\mu_{W,Y}^{3,1} - \phi\mu_\varepsilon^4 - 3[(\phi+\delta)+\phi]\mu_U^2\mu_\varepsilon^2}{(\phi+\delta)} \\
\mu_{W,Y}^{2,2} &= (\phi+\delta)\mu_{W,Y}^{3,1} - \phi(\phi+\delta)\mu_\varepsilon^4 - 3(\phi+\delta)[(\phi+\delta)+\phi]\mu_U^2\mu_\varepsilon^2 + \phi^2\mu_\varepsilon^4 \\
&\quad + [\phi^2+4(\phi+\delta)\phi+(\phi+\delta)^2]\mu_U^2\mu_\varepsilon^2 + \mu_U^2\mu_\eta^2 + \mu_\varepsilon^2\mu_\eta^2, \text{ and} \\
\mu_{W,Y}^{1,3} &= (\phi+\delta)^2\mu_{W,Y}^{3,1} - \phi(\phi+\delta)^2\mu_\varepsilon^4 - 3(\phi+\delta)^2[(\phi+\delta)+\phi]\mu_U^2\mu_\varepsilon^2 + \phi^3\mu_\varepsilon^4 \\
&\quad + 3(\phi+\delta)\mu_U^2\mu_\eta^2 + 3\phi(\phi+\delta)(\phi+(\phi+\delta))\mu_U^2\mu_\varepsilon^2 + 3\phi\mu_\varepsilon^2\mu_\eta^2,
\end{aligned}$$

so that

$$\begin{aligned}
\mu_{W,Y}^{2,2} &= (\phi+\delta)\mu_{W,Y}^{3,1} + \phi(\phi-(\phi+\delta))\mu_\varepsilon^4 + (\phi-(\phi+\delta))(\phi+2(\phi+\delta))\mu_U^2\mu_\varepsilon^2 + \mu_U^2\mu_\eta^2 + \mu_\varepsilon^2\mu_\eta^2, \text{ and} \\
\mu_{W,Y}^{1,3} &= (\phi+\delta)^2\mu_{W,Y}^{3,1} + \phi(\phi-(\phi+\delta))(\phi+(\phi+\delta))\mu_\varepsilon^4 \\
&\quad + 3(\phi+\delta)(\phi+(\phi+\delta))(\phi-(\phi+\delta))\mu_U^2\mu_\varepsilon^2 + 3(\phi+\delta)\mu_U^2\mu_\eta^2 + 3\phi\mu_\varepsilon^2\mu_\eta^2.
\end{aligned}$$

where we make use of

$$\begin{aligned}
&\phi^2 + 4(\phi+\delta)\phi + (\phi+\delta)^2 - 3(\phi+\delta)[(\phi+\delta)+\phi] \\
=\;& \phi^2 + (\phi+\delta)\phi - 2(\phi+\delta)^2 = \phi^2 - (\phi+\delta)^2 + (\phi+\delta)\phi - (\phi+\delta)^2 \\
=\;& (\phi-(\phi+\delta))(\phi+(\phi+\delta)) + (\phi+\delta)(\phi-(\phi+\delta)) \\
=\;& (\phi-(\phi+\delta))(\phi+2(\phi+\delta)),
\end{aligned}$$

and

$$\phi^3 - \phi(\phi+\delta)^2 = \phi(\phi^2-(\phi+\delta)^2) = \phi(\phi-(\phi+\delta))(\phi+(\phi+\delta))$$

and

$$\begin{aligned}
&-3(\phi+\delta)^2[(\phi+\delta)+\phi] + 3\phi(\phi+\delta)(\phi+(\phi+\delta)) \\
=\;& ((\phi+\delta)+\phi)[-3(\phi+\delta)^2 + 3\phi(\phi+\delta)] \\
=\;& 3((\phi+\delta)+\phi)(\phi+\delta)(\phi-(\phi+\delta)).
\end{aligned}$$

Given ϕ and δ are nonzero, It follows that

$$\begin{aligned}
\mu_\varepsilon^4 &= M_\varepsilon^4(\phi) \equiv \frac{1}{\phi(\phi - M_{\phi+\delta}(\phi))} \times [\mu_{W,Y}^{2,2} - M_{\phi+\delta}(\phi)\mu_{W,Y}^{3,1} \\
&\quad -(\phi - M_{\phi+\delta}(\phi))(\phi + 2M_{\phi+\delta}(\phi))M_U^2(\phi)M_\varepsilon^2(\phi) - M_U^2(\phi)M_\eta^2(\phi) - M_\varepsilon^2(\phi)M_\eta^2(\phi)]
\end{aligned}$$

and substituting for μ_ε^4 in the $\mu_{W,Y}^{1,3}$ equation gives

$$\begin{aligned}
\mu_{W,Y}^{1,3} &= (\phi+\delta)^2\mu_{W,Y}^{3,1} + \\
&\quad (\phi + (\phi+\delta))[\mu_{W,Y}^{2,2} - (\phi+\delta)\mu_{W,Y}^{3,1} - (\phi - (\phi+\delta))(\phi + 2(\phi+\delta))\mu_U^2\mu_\varepsilon^2 - \mu_U^2\mu_\eta^2 - \mu_\varepsilon^2\mu_\eta^2] \\
&\quad +3(\phi+\delta)(\phi + (\phi+\delta))(\phi - (\phi+\delta))\mu_U^2\mu_\varepsilon^2 + 3(\phi+\delta)\mu_U^2\mu_\eta^2 + 3\phi\mu_\varepsilon^2\mu_\eta^2 \\
&= (\phi+\delta)^2\mu_{W,Y}^{3,1} + (\phi + (\phi+\delta))\mu_{W,Y}^{2,2} - (\phi + (\phi+\delta))(\phi+\delta)\mu_{W,Y}^{3,1} \\
&\quad -(\phi + (\phi+\delta))(\phi - (\phi+\delta))(\phi + 2(\phi+\delta))\mu_U^2\mu_\varepsilon^2 - (\phi + (\phi+\delta))\mu_U^2\mu_\eta^2 - (\phi + (\phi+\delta))\mu_\varepsilon^2\mu_\eta^2 \\
&\quad +3(\phi+\delta)(\phi + (\phi+\delta))(\phi - (\phi+\delta))\mu_U^2\mu_\varepsilon^2 + 3(\phi+\delta)\mu_U^2\mu_\eta^2 + 3\phi\mu_\varepsilon^2\mu_\eta^2 \\
&= -\phi(\phi+\delta)\mu_{W,Y}^{3,1} + (\phi + (\phi+\delta))\mu_{W,Y}^{2,2} - (\phi + (\phi+\delta))(\phi - (\phi+\delta))^2\mu_U^2\mu_\varepsilon^2 \\
&\quad +(2(\phi+\delta) - \phi)\mu_U^2\mu_\eta^2 + (2\phi - (\phi+\delta))\mu_\varepsilon^2\mu_\eta^2,
\end{aligned}$$

where we make use of

$$\begin{aligned}
&+3(\phi+\delta)(\phi + (\phi+\delta))(\phi - (\phi+\delta)) - (\phi + (\phi+\delta))(\phi - (\phi+\delta))(\phi + 2(\phi+\delta)) \\
= \ &(\phi + (\phi+\delta))(\phi - (\phi+\delta))(3(\phi+\delta) - \phi - 2(\phi+\delta)) = -(\phi + (\phi+\delta))(\phi - (\phi+\delta))^2.
\end{aligned}$$

We therefore obtain a nonlinear equation in ϕ:

$$\begin{aligned}
\mu_{W,Y}^{1,3} &= -\phi M_{\phi+\delta}(\phi)\mu_{W,Y}^{3,1} + (\phi + M_{\phi+\delta}(\phi))\mu_{W,Y}^{2,2} - (\phi + M_{\phi+\delta}(\phi))(\phi - M_{\phi+\delta}(\phi))^2 M_U^2(\phi)M_\varepsilon^2(\phi) \\
&\quad +(2M_{\phi+\delta}(\phi) - \phi)M_U^2(\phi)M_\eta^2(\phi) + (2\phi - M_{\phi+\delta}(\phi))M_\varepsilon^2(\phi)M_\eta^2(\phi).
\end{aligned} \quad (5)$$

We have that, assuming $\phi\mu^{2,1}_{W,Y} \neq \mu^{1,2}_{W,Y}$ and $\phi\mu^3_W \neq \mu^{2,1}_{W,Y}$:

$$\begin{aligned}
M_{\phi+\delta}(\phi) &= \frac{\mu^{1,2}_{W,Y} - \phi^2 M^3_\varepsilon(\phi)}{\mu^{2,1}_{W,Y} - \phi M^3_\varepsilon(\phi)} = \frac{\mu^{1,2}_{W,Y} - \phi^2 \frac{\mu^3_W\mu^{1,2}_{W,Y} - (\mu^{2,1}_{W,Y})^2}{\phi^2\mu^3_W + \mu^{1,2}_{W,Y} - 2\phi\mu^{2,1}_{W,Y}}}{\mu^{2,1}_{W,Y} - \phi \frac{\mu^3_W\mu^{1,2}_{W,Y} - (\mu^{2,1}_{W,Y})^2}{\phi^2\mu^3_W + \mu^{1,2}_{W,Y} - 2\phi\mu^{2,1}_{W,Y}}} \\
&= \frac{\mu^{1,2}_{W,Y}(\phi^2\mu^3_W + \mu^{1,2}_{W,Y} - 2\phi\mu^{2,1}_{W,Y}) - \phi^2[\mu^3_W\mu^{1,2}_{W,Y} - (\mu^{2,1}_{W,Y})^2]}{\mu^{2,1}_{W,Y}(\phi^2\mu^3_W + \mu^{1,2}_{W,Y} - 2\phi\mu^{2,1}_{W,Y}) - \phi[\mu^3_W\mu^{1,2}_{W,Y} - (\mu^{2,1}_{W,Y})^2]} \\
&= \frac{\phi^2(\mu^{2,1}_{W,Y})^2 - 2\phi\mu^{2,1}_{W,Y}\mu^{1,2}_{W,Y} + (\mu^{1,2}_{W,Y})^2}{\phi^2\mu^{2,1}_{W,Y}\mu^3_W + \mu^{1,2}_{W,Y}\mu^{2,1}_{W,Y} - 2\phi(\mu^{2,1}_{W,Y})^2 - \phi\mu^3_W\mu^{1,2}_{W,Y} + \phi(\mu^{2,1}_{W,Y})^2} \\
&= \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})^2}{\phi\mu^3_W(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y}) - \mu^{2,1}_{W,Y}(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})} \\
&= \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})^2}{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})(\phi\mu^3_W - \mu^{2,1}_{W,Y})} = \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}.
\end{aligned}$$

Further, we have that

$$M^2_\varepsilon(\phi) = \frac{\mu_{W,Y} - M_{\phi+\delta}(\phi)\mu^2_W}{\phi - M_{\phi+\delta}(\phi)} = \frac{\mu_{W,Y} - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}\mu^2_W}{\phi - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}} = \frac{\mu_{W,Y}(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})\mu^2_W}{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}$$

and

$$\begin{aligned}
M^2_U(\phi) &\equiv \mu^2_W - M^2_\varepsilon(\phi) = M^2_U(\phi) \equiv \mu^2_W - \frac{\mu_{W,Y}(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})\mu^2_W}{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})} \\
&= \frac{(\phi\mu^3_W - \mu^{2,1}_{W,Y})(\phi\mu^2_W - \mu_{W,Y})}{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})},
\end{aligned}$$

and

$$M^2_\eta(\phi) = \mu^2_Y - M_{\phi+\delta}(\phi)^2 M^2_U(\phi) - \phi^2 M^2_\varepsilon(\phi)$$

where

$$\begin{aligned}
& -M_{\phi+\delta}(\phi)^2 M^2_U(\phi) - \phi^2 M^2_\varepsilon(\phi) \\
&= (-M_{\phi+\delta}(\phi)^2 + \phi^2) M^2_U(\phi) - \phi^2\mu^2_W \\
&= (\phi - M_{\phi+\delta}(\phi))(\phi + M_{\phi+\delta}(\phi)) M^2_U(\phi) - \phi^2\mu^2_W \\
&= \frac{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})} \frac{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) + (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})} \\
& \quad \times \frac{(\phi\mu^3_W - \mu^{2,1}_{W,Y})(\phi\mu^2_W - \mu_{W,Y})}{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})} - \phi^2\mu^2_W \\
&= (\phi\mu^2_W - \mu_{W,Y}) \frac{\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) + (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})} - \phi^2\mu^2_W
\end{aligned}$$

so that

$$M_\eta^2(\phi) = \mu_Y^2 + (\phi\mu_W^2 - \mu_{W,Y})(\phi + \frac{(\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})}) - \phi^2\mu_W^2.$$

Next, we examine the terms in (5). We have

$$\begin{aligned}
& (2M_{\phi+\delta}(\phi) - \phi)M_U^2(\phi) + (2\phi - M_{\phi+\delta}(\phi))M_\varepsilon^2(\phi) \\
= \; & (2M_{\phi+\delta}(\phi) - \phi)M_U^2(\phi) + (2\phi - M_{\phi+\delta}(\phi))(\mu_W^2 - M_U^2(\phi)) \\
= \; & 3(M_{\phi+\delta}(\phi) - \phi)M_U^2(\phi) + (2\phi - M_{\phi+\delta}(\phi))\mu_W^2 \\
= \; & -3\frac{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})}\frac{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})(\phi\mu_W^2 - \mu_{W,Y})}{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})} \\
& +(2\phi - \frac{(\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})})\mu_W^2 \\
= \; & -3(\phi\mu_W^2 - \mu_{W,Y}) + (2\phi - \frac{(\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})})\mu_W^2.
\end{aligned}$$

Further, we have that

$$\begin{aligned}
(\phi - M_{\phi+\delta}(\phi))M_U^2(\phi) &= \frac{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})}\frac{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})(\phi\mu_W^2 - \mu_{W,Y})}{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})} \\
&= (\phi\mu_W^2 - \mu_{W,Y})
\end{aligned}$$

and

$$\begin{aligned}
(\phi - M_{\phi+\delta}(\phi))M_\varepsilon^2(\phi) &= \frac{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})}\frac{\mu_{W,Y}(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})\mu_W^2}{\phi(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})} \\
&= \frac{\mu_{W,Y}(\phi\mu_W^3 - \mu_{W,Y}^{2,1}) - (\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})\mu_W^2}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})} = \mu_{W,Y} - \frac{(\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})}\mu_W^2,
\end{aligned}$$

so that

$$\begin{aligned}
& (\phi + M_{\phi+\delta}(\phi))(\phi - M_{\phi+\delta}(\phi))^2 M_U^2(\phi)M_\varepsilon^2(\phi) \\
= \; & (\phi + \frac{(\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})})(\phi\mu_W^2 - \mu_{W,Y})[\mu_{W,Y} - \frac{(\phi\mu_{W,Y}^{2,1} - \mu_{W,Y}^{1,2})}{(\phi\mu_W^3 - \mu_{W,Y}^{2,1})}\mu_W^2].
\end{aligned}$$

Thus we have that

$$\begin{aligned}
& -\mu^{1,3}_{W,Y} - \phi M_{\phi+\delta}(\phi)\mu^{3,1}_{W,Y} + (\phi + M_{\phi+\delta}(\phi))\mu^{2,2}_{W,Y} - (\phi + M_{\phi+\delta}(\phi))(\phi - M_{\phi+\delta}(\phi))^2 M^2_U(\phi) M^2_\varepsilon(\phi) \\
& +(2M_{\phi+\delta}(\phi) - \phi)M^2_U(\phi)M^2_\eta(\phi) + (2\phi - M_{\phi+\delta}(\phi))M^2_\varepsilon(\phi)M^2_\eta(\phi) \\
= & -\mu^{1,3}_{W,Y} - \phi\frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}\mu^{3,1}_{W,Y} + (\phi + \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})\mu^{2,2}_{W,Y} \\
& -(\phi + \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})(\phi\mu^2_W - \mu_{W,Y})[\mu_{W,Y} - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}\mu^2_W] \\
& +[-3(\phi\mu^2_W - \mu_{W,Y}) + (2\phi - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})\mu^2_W] \\
& \times[\mu^2_Y + (\phi\mu^2_W - \mu_{W,Y})(\phi + \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}) - \phi^2\mu^2_W] \\
= & -\mu^{1,3}_{W,Y} - \phi\frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}\mu^{3,1}_{W,Y} + (\phi + \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})\mu^{2,2}_{W,Y} \\
& +(\phi + \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})(\phi\mu^2_W - \mu_{W,Y}) \\
& \times[-(\mu_{W,Y} - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}\mu^2_W) - 3(\phi\mu^2_W - \mu_{W,Y}) + (2\phi - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})\mu^2_W] \\
& +[-3(\phi\mu^2_W - \mu_{W,Y}) + (2\phi - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})\mu^2_W][\mu^2_Y - \phi^2\mu^2_W] \\
= & -\mu^{1,3}_{W,Y} - \phi\frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})}\mu^{3,1}_{W,Y} + (\phi + \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})[\mu^{2,2}_{W,Y} + (\phi\mu^2_W - \mu_{W,Y})(-\phi\mu^2_W + 2\mu_{W,Y})] \\
& +[-3(\phi\mu^2_W - \mu_{W,Y}) + (2\phi - \frac{(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})}{(\phi\mu^3_W - \mu^{2,1}_{W,Y})})\mu^2_W][\mu^2_Y - \phi^2\mu^2_W] \\
= & \ 0.
\end{aligned}$$

Multiplying by $(\phi\mu^3_W - \mu^{2,1}_{W,Y})$ gives

$$\begin{aligned}
& -\mu^{1,3}_{W,Y}(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - \phi(\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})\mu^{3,1}_{W,Y} \\
& +[\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) + (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y})][\mu^{2,2}_{W,Y} + (\phi\mu^2_W - \mu_{W,Y})(-\phi\mu^2_W + 2\mu_{W,Y})] \\
& +[-3(\phi\mu^2_W - \mu_{W,Y})(\phi\mu^3_W - \mu^{2,1}_{W,Y}) + (2\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) - (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y}))\mu^2_W][\mu^2_Y - \phi^2\mu^2_W] = 0.
\end{aligned}$$

Note that

$$\phi(\phi\mu^3_W - \mu^{2,1}_{W,Y}) + (\phi\mu^{2,1}_{W,Y} - \mu^{1,2}_{W,Y}) = \phi^2\mu^3_W - \mu^{1,2}_{W,Y}$$

and

$$
\begin{aligned}
& -3(\phi\mu_W^2-\mu_{W,Y})(\phi\mu_W^3-\mu_{W,Y}^{2,1})+(2\phi(\phi\mu_W^3-\mu_{W,Y}^{2,1})-(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2}))\mu_W^2 \\
= \;& -3\phi^2\mu_W^2\mu_W^3+3\phi\mu_W^2\mu_{W,Y}^{2,1}+3\mu_{W,Y}\phi\mu_W^3-3\mu_{W,Y}\mu_{W,Y}^{2,1}+(2\phi^2\mu_W^3-3\phi\mu_{W,Y}^{2,1}+\mu_{W,Y}^{1,2})\mu_W^2 \\
= \;& -\phi^2\mu_W^2\mu_W^3+3\phi\mu_{W,Y}\mu_W^3-3\mu_{W,Y}\mu_{W,Y}^{2,1}+\mu_{W,Y}^{1,2}\mu_W^2 \\
= \;& -\mu_W^2(\phi^2\mu_W^3-\mu_{W,Y}^{1,2})+3\mu_{W,Y}(\phi\mu_W^3-\mu_{W,Y}^{2,1})
\end{aligned}
$$

We can then rewrite the equation as

$$
\begin{aligned}
& -\mu_{W,Y}^{1,3}(\phi\mu_W^3-\mu_{W,Y}^{2,1})-\phi(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2})\mu_{W,Y}^{3,1} \\
& +(\phi^2\mu_W^3-\mu_{W,Y}^{1,2})[\mu_{W,Y}^{2,2}+(\phi\mu_W^2-\mu_{W,Y})(-\phi\mu_W^2+2\mu_{W,Y})] \\
& +[-\mu_W^2(\phi^2\mu_W^3-\mu_{W,Y}^{1,2})+3\mu_{W,Y}(\phi\mu_W^3-\mu_{W,Y}^{2,1})][\mu_Y^2-\phi^2\mu_W^2] \\
= \;& (\phi\mu_W^3-\mu_{W,Y}^{2,1})(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}(\mu_Y^2-\phi^2\mu_W^2))-\phi(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2})\mu_{W,Y}^{3,1} \\
& +(\phi^2\mu_W^3-\mu_{W,Y}^{1,2})[\mu_{W,Y}^{2,2}+(\phi\mu_W^2-\mu_{W,Y})(-\phi\mu_W^2+2\mu_{W,Y})-\mu_W^2(\mu_Y^2-\phi^2\mu_W^2)] \\
= \;& (\phi\mu_W^3-\mu_{W,Y}^{2,1})(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}(\mu_Y^2-\phi^2\mu_W^2))-\phi(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2})\mu_{W,Y}^{3,1} \\
& +(\phi^2\mu_W^3-\mu_{W,Y}^{1,2})[\mu_{W,Y}^{2,2}+3\phi\mu_W^2\mu_{W,Y}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2] \\
= \;& \phi\mu_W^3(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}\mu_Y^2-3\mu_{W,Y}\phi^2\mu_W^2)-\mu_{W,Y}^{2,1}(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}(\mu_Y^2-\phi^2\mu_W^2)) \\
& -\phi(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2})\mu_{W,Y}^{3,1}+(\phi^2\mu_W^3-\mu_{W,Y}^{1,2})[\mu_{W,Y}^{2,2}+3\phi\mu_W^2\mu_{W,Y}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2] \\
= \;& \phi\mu_W^3(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}\mu_Y^2)-\mu_{W,Y}^{2,1}(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}(\mu_Y^2-\phi^2\mu_W^2)) \\
& -\phi(\phi\mu_{W,Y}^{2,1}-\mu_{W,Y}^{1,2})\mu_{W,Y}^{3,1}+\phi^2\mu_W^3(\mu_{W,Y}^{2,2}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2) \\
& -\mu_{W,Y}^{1,2}[\mu_{W,Y}^{2,2}+3\phi\mu_W^2\mu_{W,Y}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2] \\
= \;& \phi^2(3\mu_{W,Y}^{2,1}\mu_{W,Y}\mu_W^2-\mu_{W,Y}^{2,1}\mu_{W,Y}^{3,1}+\mu_W^3(\mu_{W,Y}^{2,2}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2)) \\
& +\phi(-\mu_{W,Y}^{1,3}\mu_W^3+3\mu_{W,Y}\mu_Y^2\mu_W^3+\mu_{W,Y}^{1,2}\mu_{W,Y}^{3,1}-3\mu_{W,Y}^{1,2}\mu_W^2\mu_{W,Y}) \\
& +\mu_{W,Y}^{2,1}\mu_{W,Y}^{1,3}-3\mu_{W,Y}^{2,1}\mu_{W,Y}\mu_Y^2-\mu_{W,Y}^{1,2}(\mu_{W,Y}^{2,2}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2) \\
= \;& 0.
\end{aligned}
$$

Thus, we have a quadratic equation of the form

$$
\begin{aligned}
& \phi^2(\mu_{W,Y}^{2,1}(3\mu_{W,Y}\mu_W^2-\mu_{W,Y}^{3,1})+\mu_W^3(\mu_{W,Y}^{2,2}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2)) \\
& +\phi(\mu_W^3(-\mu_{W,Y}^{1,3}+3\mu_{W,Y}\mu_Y^2)+\mu_{W,Y}^{1,2}(\mu_{W,Y}^{3,1}-3\mu_W^2\mu_{W,Y})) \\
& +\mu_{W,Y}^{2,1}(\mu_{W,Y}^{1,3}-3\mu_{W,Y}\mu_Y^2)-\mu_{W,Y}^{1,2}(\mu_{W,Y}^{2,2}-2\mu_{W,Y}\mu_{W,Y}-\mu_W^2\mu_Y^2) \\
= \;& A\phi^2+B\phi+C=0.
\end{aligned}
$$

The roots of this equation are given by

$$f_1 = \frac{-B - \sqrt{\Delta}}{2A.} \quad \text{and} \quad f_2 = \frac{-B + \sqrt{\Delta}}{2A.}$$

where

$$\begin{aligned}
\Delta &= B^2 - 4AC \\
&= (\mu_W^3(-\mu_{W,Y}^{1,3} + 3\mu_{W,Y}\mu_Y^2))^2 + (\mu_{W,Y}^{1,2}(\mu_{W,Y}^{3,1} - 3\mu_W^2\mu_{W,Y}))^2 \\
&\quad +2(\mu_W^3(-\mu_{W,Y}^{1,3} + 3\mu_{W,Y}\mu_Y^2))(\mu_{W,Y}^{1,2}(\mu_{W,Y}^{3,1} - 3\mu_W^2\mu_{W,Y})) \\
&\quad -4(\mu_{W,Y}^{2,1}(3\mu_{W,Y}\mu_W^2 - \mu_{W,Y}^{3,1}) + \mu_W^3(\mu_{W,Y}^{2,2} - 2\mu_{W,Y}\mu_{W,Y} - \mu_W^2\mu_Y^2)) \\
&\quad \times(\mu_{W,Y}^{2,1}(\mu_{W,Y}^{1,3} - 3\mu_{W,Y}\mu_Y^2) - \mu_{W,Y}^{1,2}(\mu_{W,Y}^{2,2} - 2\mu_{W,Y}\mu_{W,Y} - \mu_W^2\mu_Y^2)) \\
&= (\mu_W^3(-\mu_{W,Y}^{1,3} + 3\mu_{W,Y}\mu_Y^2))^2 + (\mu_{W,Y}^{1,2}(\mu_{W,Y}^{3,1} - 3\mu_W^2\mu_{W,Y}))^2 \\
&\quad +2(\mu_W^3(-\mu_{W,Y}^{1,3} + 3\mu_{W,Y}\mu_Y^2))(\mu_{W,Y}^{1,2}(\mu_{W,Y}^{3,1} - 3\mu_W^2\mu_{W,Y})) \\
&\quad -4(\mu_{W,Y}^{2,1}(3\mu_{W,Y}\mu_W^2 - \mu_{W,Y}^{3,1}) + \mu_W^3(\mu_{W,Y}^{2,2} - 2\mu_{W,Y}\mu_{W,Y} - \mu_W^2\mu_Y^2)) \\
&\quad \times(\mu_{W,Y}^{2,1}(\mu_{W,Y}^{1,3} - 3\mu_{W,Y}\mu_Y^2) - \mu_{W,Y}^{1,2}(\mu_{W,Y}^{2,2} - 2\mu_{W,Y}\mu_{W,Y} - \mu_W^2\mu_Y^2))
\end{aligned}$$

The system is symmetric in (ϕ, ε) and $(\phi + \delta, U)$. In particular, for any given solution, there exits a second solution that interchanges the role of the variables ε and U, as well as their coefficients ϕ and $\phi+\delta$. It follows that $f_1 - f_2$ corresponds to either $\phi-(\phi+\delta) = -\delta$ or $(\phi+\delta) - \phi = \delta$. (Alternatively, $M_\delta(f_1) = -M_\delta(f_2)$). Therefore, knowing the sign of δ would identify the unique root. Further, we have that $M_U^2(f_1) = M_\varepsilon^2(f_2)$ and $M_U^2(f_2) = M_\varepsilon^2(f_1)$. Therefore, knowing that $\mu_\varepsilon^2 \leq \mu_U^2$ would identify the unique root.