1. Consumer Surplus

(a) The plug-in estimator is \( \hat{A} = -\hat{\alpha}^2 / 2\hat{\beta} \).

(b) Notice

\[
\alpha = \frac{\partial}{\partial \theta} A(\theta) = \begin{pmatrix}
\frac{\partial}{\partial \alpha} (-\alpha^2 / 2\beta) \\
\frac{\partial}{\partial \beta} (-\alpha^2 / 2\beta)
\end{pmatrix}
= \begin{pmatrix}
-\alpha / \beta \\
\alpha^2 / 2\beta^2
\end{pmatrix}
\]

Thus by the Delta method, since \( \beta \neq 0 \)

\[
\sqrt{n}(\hat{\theta} - \theta) \rightarrow_d N(0, \alpha^2 V \alpha)
\]

An estimate of the asymptotic variance is \( \hat{\alpha}^2 V \hat{\alpha} \) where

\[
\hat{\alpha} = \begin{pmatrix}
-\hat{\alpha} / \hat{\beta} \\
\hat{\alpha}^2 / 2\hat{\beta}^2
\end{pmatrix}
\]

Thus a 95% confidence interval for \( A \) is

\[
\hat{A} \pm 1.96 \* \sqrt{\hat{\alpha}^2 V \hat{\alpha} / n}
\]

(c) For a bootstrap percentile interval: For \( b = 1, \ldots, B \), where \( B = 1000 \) or higher

i. Sample \( n \) iid observations from the empirical sample

ii. Construct the OLS estimates \( \hat{\theta}_b^* = (\hat{\alpha}_b^*, \hat{\beta}_b^*) \)

iii. Form \( \hat{A}^*_b = -\hat{\alpha}_b^2 / 2\hat{\beta}_b^2 \)

Given \( \{\hat{A}_1^*, \ldots, \hat{A}_B^*\} \), calculate the \( \alpha / 2 \) and \( 1 - \alpha / 2 \) empirical quantiles \( q_{\alpha / 2} \) and \( q_{1-\alpha / 2} \)

The percentile interval is \( [q_{\alpha / 2}, q_{1-\alpha / 2}] \).

2. Structural quadratic equation

(a) \( x_i^2 \) is endogenous. (If \( x_i \) and \( e_i \) are correlated, we should in general expect functions of \( x_i \) and \( e_i \) to be possibly correlated.)

(b) The number of rhs structural variables is 3, the number of instruments are 3, so this is just-identified (and thus sufficient).

(c) Given equation (2) and squaring

\[
x_i^2 = \gamma_0^2 + \gamma_1^2 z_i^2 + u_i^2 + 2\gamma_0 \gamma_1 z_i + 2\gamma_0 u_i + 2\gamma_1 z_i u_i
\]

\[
= (\gamma_0^2 + E u_i^2) + 2\gamma_0 \gamma_1 z_i + \gamma_1^2 z_i^2 + (u_i^2 - E u_i^2) + 2\gamma_0 u_i + 2\gamma_1 z_i u_i
\]

\[
= \gamma_0^2 + 2\gamma_0 \gamma_1 z_i + \gamma_1^2 z_i^2 + u_i^2
\]
say, where the error \( u_{2i} \) is mean zero (conditionally on \( z_i \)). This is the reduced form equation for \( x_i^2 \). A minimal condition for identification is that the coefficients on the instruments must be non-zero, which requires that \( \gamma_1 \neq 0 \). (If \( \gamma_1 = 0 \) then both coefficients are zero.) Furthermore, since the equation for \( x_i \) is linear in \( z_i \) it will be furthermore necessary for \( z_i^2 \) to be relevant in this equation, e.g. that the coefficient \( \gamma_1^2 \) is non-zero. However, this occurs if \( \gamma_1 \neq 0 \) so this is not an additional requirement.

In summary, \( \gamma_1 \neq 0 \) is sufficient for identification.

3. Two-step model

The question should have stated that \( E(z_i e_i) = 0 \) and \( E(z_i u_i) = 0 \).

(a) The first-step is \( \hat{\gamma} = \sum z_i x_i / \sum z_i^2 \). The second step is

\[
\hat{\beta} = \frac{\sum \hat{x}_{i2} y_i}{\sum \hat{x}_i^4} = \frac{\sum (\gamma z_i)^2 y_i}{\sum (\gamma z_i)^4} = \frac{\sum z_i^2 y_i}{\gamma^2 \sum z_i^4}
\]

(b) By the WLLN,

\[
\hat{\gamma} = \frac{1}{n} \sum z_i x_i \rightarrow_p \gamma = \frac{E z_i x_i}{E z_i^2}.
\]

Using \( y_i = \beta x_i^2 + e_i \), and then the WLLN

\[
\frac{\sum z_i^2 y_i}{\sum z_i^4} = \frac{1}{n} \sum z_i^2 x_i^2 + \frac{1}{n} \sum z_i^2 e_i
\]

\[
\rightarrow_p \beta \frac{E(z_i^2 x_i^2)}{E z_i^4} + \frac{E(z_i^2 e_i)}{E z_i^4}.
\]

since \( E z_i^2 e_i = 0 \) under \( E(e_i \mid z_i) = 0 \). (If we do not assume \( E(e_i \mid z_i) = 0 \) then there will be an extra term as well.)

One could stop here and state that

\[
\hat{\beta} \rightarrow_p \beta \frac{E(z_i^2 x_i^2)}{\gamma^2 E z_i^4} + \frac{E(z_i^2 e_i)}{\gamma^2 E z_i^4}
\]

However, we can try and say a little more. Since \( x_i = \gamma z_i + u_i \) then \( x_i^2 = \gamma^2 z_i^2 + 2 \gamma z_i u_i + u_i^2 \) so

\[
\frac{E(z_i^2 x_i^2)}{\gamma^2 E z_i^4} = 1 + 2 \frac{E(z_i^2 u_i)}{\gamma E z_i^4} + \frac{E(z_i^2 u_i^2)}{\gamma^2 E z_i^4}
\]

Thus we have

\[
\hat{\beta} \rightarrow_p \beta \left( 1 + \frac{E(z_i^3 u_i)}{\gamma E z_i^4} + \frac{E(z_i^2 u_i^2)}{\gamma^2 E z_i^4} \right) + \frac{E(z_i^2 e_i)}{\gamma^2 E z_i^4}
\]

Two of the terms drop out under slightly stronger assumptions. Suppose the linear reduced form for \( x_i \) is the conditional mean, then \( E(u_i \mid z_i) = 0 \) and thus \( E(z_i^3 u_i) = 0 \).
Furthermore if $z_i$ is exogenous in the strong sense that $E(e_i \mid z_i) = 0$ then $E(z_i^2 e_i) = 0$ as well. Under these two plausible restrictions we find

$$\hat{\beta} \rightarrow_p \beta \left( 1 + \frac{E(z_i^2 u_i^2)}{\gamma^2 E z_i^4} \right).$$

Finally if we make the even stronger assumption that the error $u_i^2$ is conditionally homoskedastic $E(u_i^2 \mid z_i) = \sigma_u^2$ then the limit simplifies to

$$\hat{\beta} \rightarrow_p \beta \left( 1 + \frac{\gamma^2 E z_i^2 u_i^2}{E z_i^4} \right) = \gamma^2 + \frac{E z_i^2 u_i^2}{E z_i^4},$$

where the second equality holds if we assume $E(u_i \mid z_i) = 0$, and the third holds if we further assume that $u_i$ is independent of $z_i$. We thus have that

$$\hat{\beta} \rightarrow_p \beta^* = \beta + \frac{\gamma^2 E z_i^2 u_i^2}{E z_i^4}.$$  \hfill (2)

(c) In general, $\hat{\beta}$ is not consistent for $\beta$, as the limit obtained in the previous sub-question is not simply $\beta$. However, we found that if $E(e_i \mid z_i) = 0$, $E(u_i \mid z_i) = 0$, and $E(u_i^2 \mid z_i) = \sigma_u^2$ then we found the simplified limit (2). Are there conditions under which this simplifies to $\beta$? For $\beta \neq 0$ the answer is no. In fact, the plim $\beta^*$ is biased away from 0, in the sense that if $\beta > 0$ then $\beta^* > \beta$ and if $\beta < 0$ then $\beta^* < \beta$. So there is no condition under which the estimator is generally consistent.

There is, however, one special condition. If $\beta = 0$ then the limit in (2) is also 0, so $\hat{\beta} \rightarrow_p 0$ and the estimator is consistent. If you go back to the more general case (1) we see that $E(e_i \mid z_i) = 0$ and $\beta = 0$ are sufficient for $\hat{\beta} \rightarrow_p 0$ and hence consistency.

It might seem somewhat odd to say “The estimator is consistent when the true coefficient is zero” but this is a general finding for two-step estimators.

4. The relation between the structural and reduced forms is $\beta = \Gamma \lambda$. Partitioning,

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \Gamma_{11} & \Gamma_{21} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} I & \Gamma_{21} \\ 0 & \Gamma_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

since $x_{1i}$ is exogenous and in both the structural and reduced form equations. Thus

$$\beta_1 = \lambda_1 + \Gamma_{21} \lambda_2$$

$$\beta_2 = \Gamma_{22} \lambda_2$$

The hypothesis is $\beta_2 = 0$, which holds if and only if $\lambda_2 = 0$ (if $\Gamma_{22}$ is full rank which holds if the model is identified). Thus the hypothesis is equivalent to the restriction $\lambda_2 = 0$, which is a restriction on the reduced form equation (3).

The hypothesis $\lambda = 0$ can be tested in equation (3) by OLS estimation of (3) and a (heteroskedasticity-
robust) Wald test for $\lambda_2 = 0$

$$W = \hat{\lambda}_2' V_{\hat{\lambda}_2}^{-1} \hat{\lambda}_2$$

Under $H_0 : \lambda_2 = 0$ the statistic converges to a $\chi^2_{t_2}$ distribution. We reject $H_0$ in favor of $H_1 : \lambda_2 \neq 0$ at the level $\alpha$ if $W > c$ where $c$ satisfies $P(\chi^2_{t_2} > c) = \alpha$. 